

Brazil's R&D agenda

The new Brazilian president has so far impressed with his handling of the country's shaky economy. He should now seize an opportunity to transform Brazil's ability to use science and technology for its economic development.

Luiz Inácio Lula da Silva alarmed many of Brazil's élite last year, as his campaign for the country's presidency cruised towards a comfortable victory. A left-wing former trade-union leader with the popular touch, Lula had a clear mandate for change — and scientists were among those who were anxious about his plans.

Almost six months into his administration, however, Lula has maintained a steady course in his handling of Brazil's economy, burdened as it is by huge international debts. He has also reassured the scientific community (see page 379). Some researchers still question the qualifications of science minister Roberto Amaral, but ministries and science agencies are being filled with senior officials who have strong credentials, and so far the federal science budget has been protected from cuts that are being forced elsewhere.

Lula is also extremely charismatic. Although he is no intellectual, having quit formal education at a relatively early age, academics and others in the scientific establishment have been impressed by his ability both to focus on essentials and to motivate others to deal with them. All the more reason, then, for Lula to move quickly to seize a key opportunity for Brazil: to turn its significant strengths in both physical and biological sciences into long-term economic advantage.

Despite severe problems with Brazil's education system and intolerable social disparities, there is no shortage of young talent keen to study science in universities, and the annual output of PhDs has risen to more than 6,000. In a huge country where individual states play as big a role as the federal government in stimulating economic development, São Paulo stands out not only for its commitment to its state university and to science funding, but also for its strategic approach to science and technology. This includes creating conditions to favour the launch of high-tech companies, attracting foreign industrial

investment, and support for research centres that will not only deliver high-quality science but also foster commercial spin-offs.

Such an approach cannot be duplicated at a national level even in the medium term, as some states have only a rudimentary science base. But Lula has stated his desire to widen access to university education and to spread research funding around the country, away from the wealthy southern states. His ministers, meanwhile, are publicly backing plans to establish new centres of excellence (see page 372). These are laudable goals, but a balance is needed: Lula will limit his country's potential if he dissipates such initiatives too widely across his territories, or lowers the thresholds of university entry too far.

Above all, Lula needs to take a personal lead in developing a national strategy for innovation and development, and should establish a better structure to assist him. A national council for science and technology is already in place, with representatives from key ministries including that of science and technology. But it meets infrequently, and its role remains to be established.

Lula needs an adviser who has scientific clout plus the insight to understand how technology not only depends on basic science but can also be stimulated and exploited to national ends — something along the lines of the director of the US Office of Science and Technology Policy, perhaps. That person should have direct access to the president and be given a key responsibility: to work with the national science and technology council and others to develop a ten-year vision by which top-down initiatives, ranging from fiscal incentives to directed research funding, can complement the bottom-up motivation that exists aplenty in Brazil's universities. If he can pick the right individual for the job, Lula could enable Brazil to become the economic powerhouse that it deserves to be. ■

Trade war: what is it good for?

The United States' challenge to Europe's ban on approving new transgenic foods could yield nothing but losers.

The transatlantic transgenic trade war has begun. Since 1998, when the European Union (EU) stopped approving new transgenic food crops for sale, US farmers have been seething. And last week, their government finally moved to open EU markets to genetically modified (GM) food, mounting a legal challenge to Europe's moratorium.

While few in US government or industry are happy about the EU's position, even fewer believe there is much to gain from taking the case to the World Trade Organization (WTO) at this late stage. The State Department, in particular, is opposed to anything that would strain already fractious US–European relations still further. Crop biotechnology companies would also have preferred a diplomatic route.

There are various motivations for the WTO suit (see page 369). But in the event, the action may have as much to do with domestic US politics as anything else. Senator Chuck Grassley (Republican, Iowa) had warned that he would speak out if the challenge were not filed soon. As chair of the Senate Committee on Finance, responsible for pushing through the administration's tax cuts, Grassley is

not someone with whom President George W. Bush wanted to fall out.

The problem is that the suit stands to do little good. The EU moratorium had appeared likely to end later this year anyway, as soon as strict rules requiring GM produce to be labelled are finalized. The danger now is that EU member states might respond by continuing their moratorium for another year or so, until the suit is resolved. Even if the WTO rules in favour of the United States, an aggrieved Europe might defy the ruling and pay any retaliatory import duties that the United States would then impose.

Such tit-for-tat measures are in no one's interest, however. Although it might be seen as capitulation, the sensible response would be for the EU to swallow its pride and lift its embargo, as soon as its labelling rules are in place. So far, there is no scientific evidence suggesting that GM crops pose a special risk to human health or the environment. Countries may still reject any GM crop or product that is subsequently shown to pose a demonstrable hazard. And once labelling is in place, consumers who distrust the technology can simply reject GM food in the marketplace. ■





Invisible enemy
Iraqi looting sparks fears over access to radioactive sources
p370



All in the mind
Neuroscientists lay plans for Brazilian research community
p372



Ethical tangle
Germany sets out to secure international cloning ban
p373



Bone home
UK museums may lose grip on human artefacts
p374

Trade war looms as US launches challenge over transgenic crops

Jonathan Knight

The opening shots in a transgenic trade war were fired last week, as the United States delivered a warning to Europe — and perhaps to developing nations as well — that trade barriers to genetically modified crops are not acceptable. But observers warn that the conflict is likely to be a messy one, and that the outcome is far from certain.

The United States opened hostilities on 13 May by announcing that it will ask the World Trade Organization (WTO) to declare that the de facto moratorium on approving new transgenic crops in the European Union (EU) is illegal. The EU's approval process stalled in 1998, when six member states called for stricter laws on labelling genetically modified food. Joining the United States in the suit are Canada, Argentina, Egypt, seven Latin American countries, New Zealand and Australia.

Some researchers say that the EU is not the only target of the move. "There is a component that is a message to the developing world," says Alan McHughen, a plant-biotechnology researcher at the University of California, Riverside. Last year, for example, Zambia rejected shipments of US food aid containing transgenic corn. Officials there claimed they were concerned that some of the corn might be planted, thereby jeopardizing the country's export market to Europe (see *Nature* **418**, 571–572; 2002). McHughen says the United States wants to show that it is dedicated to opening up the European market, and that trade sanctions will not be tolerated anywhere.

But the WTO case may not be enough to achieve that aim, say food-policy experts. The EU is expected to pass new rules later this year requiring stricter labelling and traceability of food containing genetically modified ingredients. Currently, only products that contain detectable amounts of transgenic DNA or protein must be labelled. The new rules would extend the requirements to all food and animal feed derived from genetically modified plants. 'Farm-to-fork' records of how all transgenic components were used at each step of food production would also have to be kept.

But these requirements could exclude many developing nations from selling trans-



All in the label: European opposition to transgenic foods could keep them off supermarket shelves.

genic products to Europe, because such countries lack the infrastructure to ensure the traceability of each boatload of grain and rice. "With relatively poor countries such as Thailand and Kenya, you are effectively shutting them out of the biotech market," says Gregory Conko, director of food-safety policy at the Competitive Enterprise Institute, a think-tank based in Washington DC.

High price

Whatever the effect on the developing world, US agribusinesses are unlikely to be exporting transgenic crops to Europe in the near future. The WTO case could take up to 18 months to resolve. And EU member states that had been expected to lift the moratorium this autumn, once the stricter labelling rules are in place, may choose to wait until the dispute is resolved.

Even if the United States wins its case, the most that the WTO can do is allow it to set up import tariffs as compensation for the US\$300 million it claims to be losing in farm exports to Europe every year. "It seems likely that it will heighten rather than lessen the dispute between the United States and Europe," says Gary Toenniessen, director of food security at the Rockefeller Foundation in New York.

If the EU does lift the moratorium, many US exporters and European retailers may still be deterred by the new rules. "If the labelling and traceability requirements are the price we have to pay to get the moratorium lifted, we are paying too much," says Robert Paarlberg, an expert on international agricultural policy at Wellesley College near Boston.

To get transgenic crops into Europe, says Conko, the United States or one of its allies will have to contest the new rules in another WTO challenge. But Michael Hansen, a food-policy advocate with the Consumers Union in New York, says that the WTO guidelines already back strict labelling, and that the United States wouldn't win a WTO challenge on that issue.

Even when the trade wars are resolved, consumer opinion could still scupper the entry of transgenic crops into the European market. Some experts, such as Conko, say that the opposition to genetically modified food seen in public-opinion surveys does not imply that consumers would shun it. But many observers say that the surveys indicate real opposition among Europeans. "Europe will not like being brought before the WTO at this time," says Paarlberg. "The United States is not popular, the WTO is not popular, and genetically modified foods are not popular." ■

Iraqi looters spark alert over radiation risks

Geoff Brumfiel

Widespread looting in Iraq has raised fears about the fate of some 1,100 radioactive sources at sites across the country.

The sources, which were used in hospitals and industrial facilities, were catalogued by the International Atomic Energy Agency (IAEA) in the mid-1990s, but their fate in the weeks since the war ended is uncertain.

"There are some large and dangerous sources among them," says Melissa Flemming, a spokeswoman for the IAEA in Vienna. "We are deeply concerned about the possibility that some of this material has been broken into."

At Iraq's largest nuclear facility, the Tuwaitha Nuclear Research Centre near Baghdad, civilians have emptied containers of concentrated uranium oxide and used the metal drums to store food and water. The theft has been followed by reports of illness, but many scientists question whether uranium is the cause. Most of the metal was only slightly radioactive and would not have induced immediate radiation sickness in those handling it, says Michael Levi, a physicist at the Brookings Institution in Washington. Civilians would also have to ingest large amounts to feel the toxic effects of the heavy metal, he adds.

But uranium was not the only source of radiation at Tuwaitha. A 1993 IAEA inspection listed a partial inventory of 177 sources stored there, including 52 sources of cobalt-60 and 32 sources of caesium-137, as well as stores of other potentially dangerous isotopes such as iridium-192, strontium-90 and radium-226. Some of those sources are "very powerful" and pose a threat to anyone who might come across them, says David Albright, a former nuclear inspector



Safe and sound? Iraq's Tuwaitha Nuclear Research Centre holds a range of highly dangerous materials.

who has visited Tuwaitha and who now directs the Institute for Science and International Security in Washington. "My understanding is that some of them could be fatal," he says.

According to Flemming, 950 similar sources of varying size and power are scattered throughout the country. They were used in applications such as radiotherapy, X-rays, welding and oil surveying, she says. Such loose sources are dangerous because they can be mistaken for scrap metal and spread throughout the community. The most notorious incident occurred in 1987 in Goiânia, Brazil, when more than 200 people, several of whom later died, were exposed to a caesium-137 source scavenged from an abandoned hospital.

Because of their potential toxicity, these sources could also be valuable material for

terrorists, says Joseph Cirincione, head of the nonproliferation project at the Carnegie Endowment for International Peace in Washington. "I'm less worried about the uranium that seems to have disappeared than the highly radioactive elements that would be perfect for enhancing an al-Qaeda truck bomb," he says. So far, there is no evidence that any material has found its way into the hands of terrorists.

The looting has increased the pressure on the United States to allow IAEA inspectors back into the country. The US army has deployed its nuclear-disablement team to assess Iraq's nuclear holdings, but Albright and others believe that the true condition of Iraq's nuclear sources will not be learned until the IAEA resumes inspections. "The United States is not on top of this," he says. ■

Animal studies hint at staying power of SARS

Helen Pearson, New York

Studies of viruses that infect farmyard and domestic animals indicate just how difficult severe acute respiratory syndrome (SARS) will be to tackle, scientists at a New York Academy of Sciences meeting heard last weekend.

Coronaviruses from the same family as that believed to cause SARS can trigger a variety of comparable lung and gut diseases in animals from chickens to cats. Factors that worsen animal infections might suggest why some SARS patients become highly infectious 'super-spreaders' or succumb more easily to the disease, delegates were told.

One tentative theory — that human patients infected with another virus or

bacterium develop more severe symptoms when they catch SARS — was backed up by Linda Saif, a coronavirus expert at Ohio State University in Wooster. In unpublished work, Saif has shown that pigs carrying a lung infection before being inoculated with porcine respiratory coronavirus suffer a worse fever and longer infection than previously healthy swine.

Some SARS patients have been found to carry a common respiratory virus called human metapneumovirus. Donald Low, a microbiologist at Mount Sinai Hospital in Toronto, agreed that co-infection with this or another pathogen might exacerbate SARS symptoms, but researchers say that there are too little data on SARS to know whether

Saif's work is relevant to humans.

Researchers also discussed the probability that patients who are no longer showing symptoms of SARS could continue to spread the disease. Some national health authorities currently release patients from isolation once their symptoms disappear. But Kathryn Holmes, who studies coronaviruses at the University of Colorado Health Sciences Center in Denver, pointed out that cats infected with a coronavirus can continue to shed the virus for up to seven months after symptoms clear up (A. A. Herrewegh *et al. Virology* 234, 349–363; 1997).

Saif also discussed research that adds weight to concerns that the steroids commonly used to treat SARS in Hong Kong

Lack of trust hampers hunt for weapons

Declan Butler

Confusion over how the US-led occupying powers will treat Iraqi scientists is hindering the search for weapons of mass destruction in the country, according to a growing number of weapons experts. Until the uncertainty ends, say observers, scientists who could hand over valuable information may be tempted to hide or flee to rogue nations.

The arrest of prominent Iraqi scientists, such as Huda Salih Mahdi Ammash, suspected of developing biological weapons, and Rihab Taha, a senior figure in the country's germ-warfare programme, have recently made the headlines. Yet evidence of Iraq's alleged programmes for biological, chemical or nuclear weapons remains elusive, and some observers suggest that the coalition's treatment of Iraqi scientists may be partly to blame.

US radio broadcasts in Iraq have asked scientists to step forward, promising that they "will be treated with respect and dignity". But David Albright, president of the Washington-based Institute for Science and International Security and a former nuclear-weapons inspector in Iraq, says that he is in daily contact with Iraqi scientists who fear being taken prisoner by US forces, retaliation by Saddam loyalists, or kidnap and ransom to coalition authorities. Albright believes that all three are happening.

Concern that they may be treated as criminals removes scientists' incentive to cooperate, says Ibrahim al-Marashi, a research associate at the Center for Non-proliferation Studies in Monterey, California. Even those in custody are unlikely to cooperate until they receive assurances that they will not be prosecuted for war crimes, adds Jonathan Tucker, a weapons expert at the US Institute of Peace in Washington.

might instead be exacerbating the disease (see *Nature* 423, 4; 2003). She warned that dexamethasone, an anti-inflammatory corticosteroid, can actually worsen a coronavirus infection in cows and prolong the period of time for which they release virus (H. Tsunemitsu, D. R. Smith and L. J. Saif *Arch. Virol.* 144, 167–175; 1999).

Saif also cautioned that researchers have had limited success in producing vaccines against animal coronaviruses. Several studies have shown that experimental vaccines against feline infectious peritonitis virus (FIPV), for example, can aggravate the infection through a phenomenon called antibody-dependent enhancement. Antibodies produced in response to



Huda Salih Mahdi Ammash is under arrest, but should rank-and-file Iraqi scientists face the same fate?

Officials at the United Nations Monitoring, Verification and Inspection Commission (UNMOVIC), which the United States has not yet allowed back into Iraq, say that the cooperation of Iraqi scientists is crucial. Without their help, say commission staff, it will be difficult to gain an overview of what weapons Iraq might have produced, and in what quantity.

Others experts add that Iraqi scientists will be able to shed light on crucial issues, such as

the fate of Iraq's dual-use materials — substances that could be used to create weapons but which could also have peaceful applications. And some observers note that, in the past, weapons programmes have remained hidden until being revealed by scientists.

Tucker argues that weapons scientists who were not key figures in Saddam's regime should be given an amnesty and incentives to cooperate, to discourage them from fleeing or selling their talents elsewhere. He recommends that a centre be set up in Baghdad to offer weapons scientists stable employment, along the lines of the International Science and Technology Center in Moscow, an inter-governmental organization established in 1992 to divert the talents of Soviet weapons scientists to peaceful ends.

But others question how much information Iraqi scientists will be able to supply. "Very few people were probably in the know — only those closest to Saddam or the 'simpletons' who did the actual concealment work," says Ephraim Asculai of the Jaffee Center for Strategic Studies at Tel Aviv University in Israel.

The current coalition weapons inspectors are due to be replaced in the coming weeks by the Iraq Survey Group, a 2,000-strong Pentagon-led body that will investigate everything from weapons and potential war crimes to terrorist connections. The coalition has not yet responded to requests to allow inspectors from UNMOVIC or the International Atomic Energy Agency to return to Iraq.

vaccination with FIPV proteins bind to the virus and boost its uptake by immune cells called macrophages. Other viruses are destroyed when engulfed by macrophages, but in this case the FIPV multiplies inside the cells.

This suggests that one possible strategy to fight SARS — collecting and injecting antibodies from recovering patients into new patients — might similarly inflame the disease. Pharmaceutical-industry representatives at the meeting reiterated that the hope of a quick drug or vaccine for SARS is slim. A new antiviral drug would take a minimum of four to five years to reach the shelf, they said, a vaccine closer to six.

♦ www.nyas.columbia.edu

Brazilian brain experts plan research village

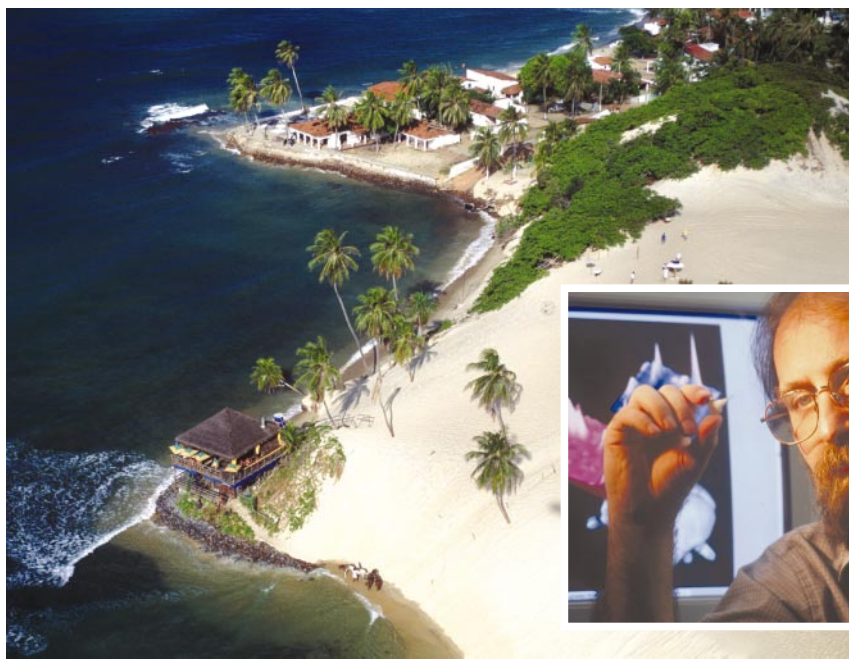
Hannah Hoag

Three expatriate Brazilian neuroscientists have launched a bid to build a research centre in the northern coastal city of Natal in Brazil. According to their plans, the institute will focus on neuroscience, and will include a mental-health clinic and a school for underprivileged children.

The move is well timed — Brazil's new president, Luiz Inácio Lula da Silva, has pledged to boost the country's research budget and promote science in areas away from the traditional centres of excellence in Rio de Janeiro and São Paulo in the south (see News Feature, page 379). But those involved in the bid admit that it will be difficult to raise the US\$20 million needed to set up the centre.

The three neuroscientists behind the plan are Miguel Nicolelis and Sidarta Ribeiro of Duke University in Durham, North Carolina, and Claudio Mello of the Oregon Health and Science University in Portland. They say that the new institute will employ expertise in molecular and cellular biology and electrophysiology in an attempt to understand the basic mechanisms that control cellular circuits within the brain.

The plans have already gained some eminent backers. Nicolelis says that Nobel laureate Torsten Wiesel, president emeritus of Rockefeller University in New York, has



Life's a beach: Miguel Nicolelis (inset) hopes to build a neuroscience campus at Natal in north Brazil.

agreed to be a trustee. And Jon Kaas, a neuroscientist at Vanderbilt University in Nashville, Tennessee, who has collaborated with Brazilian researchers for two decades, will also join the board. "There is a lot of very

good talent in Brazil, but they have had a hard time getting research time and research facilities," says Kaas.

Additional facilities such as the mental-health centre and the children's school will dot the institute's campus. "There will be a large social programme linked to the scientific endeavour," says Nicolelis. The school, for example, will enrol over 300 underprivileged toddlers and provide them with an education rich in science, language and arts until their late teens.

But funding the institute is likely to be a struggle. The neuroscientists plan to collect money through a non-profit organization which will be named after the Brazilian aviation pioneer, Alberto Santos-Dumont. Nicolelis estimates that they will need over \$20 million to build the institute and set up an endowment. But he has received only \$50,000 in pledges from private donors, and he concedes that most of this will be used to finance an international symposium on neuroscience scheduled to take place in Natal in March 2004.

Brazilian politicians, including science minister Roberto Amaral and education minister Cristovam Buarque, have backed the project. But officials at the Brazilian government say that their contribution will primarily be a donation of a 150 acres of land. The majority of funding for Brazilian science currently comes from state and federal government, raising questions about how much money Nicolelis and his colleagues will be able to raise from private sources in the country.

Spiralling costs dog comet mission

Declan Butler, Paris

Europe's troubled Rosetta comet mission has a new target, but it is not out of the woods yet — a lack of funds still threatens the project.

The Rosetta craft, which is designed to land on a comet, was scheduled for lift-off in January this year, but the launch was postponed as a result of technical problems with the new class of Ariane rockets used by the European Space Agency (ESA). The delay meant that the probe missed the window to reach its original target, the comet Wirtanen (see *Nature* 421, 301–302; 2003).

At a meeting of ESA's science committee last week, officials confirmed that the new target will be the comet Churyumov-Gerasimenko. But before the launch, which will take place on 26 February 2004 on a standard Ariane 5 rocket, extra money must be raised. The cost of storing the satellite and redesigning the mission will be some €70 million (US\$82 million).

David Southwood, ESA's head of science, says that nobody wants Rosetta to be cancelled, especially as most of its costs have already been paid. But ESA's science

programme is facing a cashflow crisis.

Problems with the Ariane rockets have added to the costs of several planned missions, and a downturn in the commercial space market means that ESA's industrial partners are now demanding more money up front for their involvement in collaborative projects.

Last week's meeting heard that ESA's science budget is facing an overrun of €140 million this year, and this is set to increase in the following two years. To launch Rosetta, the science programme will need a loan from ESA's governing body.

ESA has also had to pay out more than €70 million to support scientists building instruments for its Herschel infrared space observatory and for its Planck mission, which will study cosmic microwaves. Both projects are slated to launch together in 2007. ESA traditionally pays spacecraft and launch costs, with member states footing the bill for instruments. But the agency is increasingly having to pay for instruments, and officials say that future payloads will be subject to contracts requiring member states to commit to full funding for the project's duration. ■

T. SVENSSON/CORBIS; J. WALLACE/DUKE PHOTO (INSET)

Copied citations give impact factors a boost

Tom Clarke

Scientific papers that are not widely read and that lack any great influence can end up being classed as high-impact, claim researchers in California.

The mistake occurs because citations are often just copied from the reference list of one paper to another. A largely unremarkable or unread paper can therefore end up becoming highly cited, the researchers suggest.

"Simple mathematical probability, not genius, can explain why some papers are cited a lot more than others," says Vwani Roychowdhury, an electrical engineer at the University of California, Los Angeles.

The assertion hinges on a previous analysis by Roychowdhury and his colleague Mikhail Simkin. Last year, they tracked identical errors in reference lists citing a seminal 1973 paper and concluded that almost 80% of authors had not read the paper in question before citing it (see *Nature* **420**, 594; 2002).

The pair have now built on that finding to generate a mathematical model to predict citation levels. They tested their prediction in an analysis of about 24,000 articles from the journal *Physical Review D* stored on



SPIRES, a database of high-energy physics papers. The database sorts papers into six categories according to the number of citations that they receive — those receiving 500 or more are classed as 'renowned'.

Roychowdhury and Simkin's model closely matched the real distribution of citations. In results posted on the arXiv preprint server, they predicted that 40 papers would be cited 500 times or more. In reality, 44

articles in *Physical Review D* are renowned.

"If people cite randomly, the citation distribution would be the same as in reality," says Roychowdhury. Given that citation patterns are similar in other sciences besides physics, the outcome of the model should be similar for biology or engineering papers, he argues.

The idea is "fascinating, unorthodox and inspiring," says bibliometrics expert Anthony van Raan of Leiden University in the Netherlands. But although van Raan's research also suggests that copying is rife, he believes that citations levels generally reflect a paper's true impact. Researchers tend to copy references rationally, rather than randomly, and do so from papers that they have read, he says.

Ben Martin, director of the Science Policy Research Unit at the University of Sussex in Brighton, UK, points out that the new model is important, even if it only partially reflects reality. "Citations are a surrogate for quality," he says. "Some departments look at these when making decisions in recruitment or tenure. Even if this is just a quirk, it's one that is worth airing."

♦ <http://xxx.lanl.gov/abs/cond-mat/0305150>

Researchers divided over ethics of a ban on cloning

Alison Abbott, Berlin

The German government last week began a push for an international ban on all types of human cloning.

But at a meeting to discuss the issue — held in Berlin and sponsored by the German science ministry — some researchers claimed that even in the usually clear-cut case of reproductive cloning, current ethical arguments in favour of a ban are far too simplistic.

The meeting was scheduled after the country's parliament voted last month to ask the United Nations to call for an international ban on both therapeutic and reproductive cloning, practices already illegal in Germany.

The ethics of therapeutic cloning, in which stem cells are produced for research purposes, are most controversial — some countries, such as Britain and China, allow the practice, whereas others, such as Italy, do not. Suggestions for bans on reproductive cloning usually attract unanimous support, with ethicists arguing that few cloned embryos result in live births, and that surviving clones are usually abnormal or sick.

But Eckhard Wolf of the University of Munich, an expert on the cloning of farm animals, is one of several at the Berlin

meeting who predicted that these safety issues will eventually be resolved. He argued that the ethical objections to a ban will then become much more difficult to resolve.

The suggestion is controversial. Other researchers present, including Rudolf Jaenisch, a geneticist at the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, asserted that some problems with cloned embryos, such as faulty gene regulation, are intrinsic to the process.

But ethicists note that if safety is taken out of the equation, other arguments against reproductive cloning — such as the psychological pressure that clones may find themselves under because their 'twin' has already lived an earlier life — do not withstand ethical analysis.

"It is hard to identify a moral or human right why a healthy cloned child should be different from a healthy normal child," said Dan Brock, a bioethicist at the US National Institutes of Health in Bethesda. "Even the concept of the right to a unique identity is not disturbed because people develop according to their circumstances, not their genes."

Ortwin Renn, a philosopher at the Academy of Technological Assessment in Stuttgart, added that in the absence of

safety concerns, controversy would centre on the time after conception that the embryo is recognized either morally or constitutionally as having equal rights with a person.

Catholics, for example, see this as occurring at the moment of conception, whereas eastern philosophies assign rights at the moment of birth. Even within a culture, it is difficult to get consensus. This, suggests Renn, is why Germany tries to defer this "fundamentally irresolvable" decision to a higher authority such as the United Nations.



Double trouble: opponents of cloning make their case in Berlin last week.

A. ABBOTT

Australian reform will see universities put squeeze on students

Sydney The Australian government has unveiled a blueprint for higher-education reform that will allow universities to ratchet up student fees and clamp down on industrial action.

The A\$1.5-billion (US\$1-billion) package, announced in the federal budget on 12 May, will increase higher-education funding as long as universities comply with a raft of reforms, such as the abolition of compulsory student unions. These measures, together with increased student fees and restrictions on industrial action, make it likely that the reform package will face a rough ride in the Senate, in which the ruling Liberal Party lacks a controlling majority.

The plan has generated mixed reactions among university staff. Funding for student courses is "disappointingly modest", says Paul Greenfield, deputy vice-chancellor of the University of Queensland in Brisbane. However, university chiefs have welcomed the partial deregulation of student fees, given the limited public funds available.

University research will not be affected by the reform, as this will be addressed later in the year when the government completes its audit of the nation's scientific activities.

Museums face prospect of giving up human remains

London Thousands of prehistoric human bones and skulls, brought to Britain during the nineteenth century, may be set to return home. A report commissioned by the British government, which will be released in the summer, will propose legal changes that would allow indigenous people to claim items from British museums and galleries.

But anthropologists last week warned that handing over the remains would be a blow to evolutionary science, particularly if they were reburied or damaged. The collections are important for understanding human evolution and migration, they say.

"Human specimens help us to understand



Bone of contention: Britain's museums may have to return human artefacts to their places of origin.

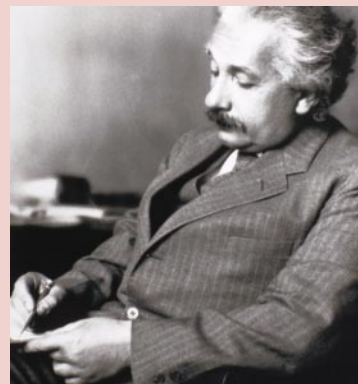
Einstein enthusiasts gain net benefit

Pasadena Albert Einstein's musings on science, politics and travel are to be made available to the public on the Internet.

More than 230 scientific manuscripts, 740 non-scientific essays and 5 travel diaries — many previously unseen by the public — have been compiled in a free, searchable database.

The project coincides with a symposium on papers by Einstein (pictured), hosted by the American Museum of Natural History in New York. "Einstein interests a broad intellectual community and he is a public figure for science," says science historian Jürgen Renn of the Max Planck Institute for the History of Science in Berlin, who proposed the project more than a decade ago.

♦ www.alberteinstein.info



patterns of health and disease in earlier populations," says Louise Humphrey, an anthropologist at the Natural History Museum in London. The museum houses a collection of around 20,000 items, many representing indigenous people from Australia and North America.

In the United States, anthropologists have been at loggerheads with native tribes since the enactment of 1990 Native American Graves Protection and Repatriation Act.

Approval for stem cells pre-empts parliament

Ottawa The Canadian Institutes of Health Research are to bypass a legislative hold-up that is stalling plans to use human embryonic stem cells for research, the institutes' president Alan Bernstein has announced.

The Canadian parliament is considering a bill that will allow researchers to work on stem cells taken from unused embryos at fertility clinics, but would outlaw cloning to produce such cells. The legislation was expected to be approved earlier this year, but its passage has been delayed by controversy over stem-cell research. The bill is still expected to become law later this year, but Bernstein says that he will begin to implement the guidelines before then, by assembling an ethical-oversight committee for the work.

Oponents of stem-cell research have accused Bernstein of circumventing the legislative process. But research-advocacy groups have welcomed the move, arguing that the parliament's indecision has held up Canada's contribution to stem-cell science.

Italian research chief jumps ship as reorganization looms

Rome Lucio Bianco last week resigned as president of Italy's main basic-research organization, the CNR. The announcement came one day before the Italian government issued a decree that places the CNR, as well as the National Institute of Astrophysics (INAF)

and the Italian Space Agency, in the hands of commissioners for reorganization.

The decree was drawn up last autumn by research minister Letizia Moratti in a bid to make research more productive. But the research community viewed it as an attempt by the government to gain political control over basic research (see *Nature* 421, 465; 2003).

The decree's exact contents are not yet known, but the National Institute for the Physics of Matter is likely to be brought into the CNR, and astronomy institutes are likely to be transferred from the CNR to the INAF. Researchers say they will consider a protest if they are not consulted over the reorganization.

A career in sport? I'd rather be on the bench

Tokyo More Japanese schoolboys want to be academics when they grow up than anything else, says a survey released this month. The survey, which asked 991 children of up to primary-school age what they wanted to be, found that an academic career was the choice for 9.6% of boys — more than wanted to play professional soccer (9.1%) or baseball (8.5%).

It is the first time that academia has taken the top spot since the Dai-ichi Mutual Life Insurance Company launched the annual survey in 1989. The company attributes the jump to the fact that Nobel Prizes went to Japanese scientists in 2000 and 2001.

Hopes for a better gender balance in future research are likely to be dashed, however. Girls in the survey were more interested in catering and nursing, with an academic career not making it into the top ten.

Correction

In the news feature "The tiny toolkit" (*Nature* 423, 10–12; 2003), an image of a cantilever on page 11 was incorrectly credited to the California Institute of Technology. It was actually developed by teams at IBM Zurich Research Laboratory and the University of Basel. On page 12, the diagram labels have been mixed up — the left-hand protein should be marked 'field off' and the right-hand protein 'field on'.



Clear as mud

It's not surprising that some academic papers seem to swim before our eyes — the scientific literature has become steadily less accessible over the past half-century. Can we stop this trend, asks Jonathan Knight.

"There is no form of prose more difficult to understand and more tedious to read than the average scientific paper," wrote Francis Crick in his 1994 book *The Astonishing Hypothesis*¹. The observation is a caution to lay readers tempted to delve into the papers referenced in the book. But the co-discoverer of the structure of DNA was also acknowledging what everyone in science knows: research papers can be a nightmare to read.

It wasn't always so. Crick and others of his generation, who began writing scientific papers in the 1940s, have witnessed the transformation of scientific prose. A form that was as readable as the average newspaper has, in some fields, become a jungle of jargon that even those familiar with the territory struggle to understand.

The balkanization of science into sub-disciplines, each with its own vocabulary, is largely to blame. Many journals are trying to tackle this, producing easy-to-read summaries of papers, and linking online papers to web-based glossaries. But these approaches tend to have a limited impact, whereas addressing other factors — notably writing style — could transform many papers. Writing takes practice, yet it is not part of standard scientific training. So could science become

readable again if researchers went back to school and took writing lessons?

Readability itself is not easy to quantify. Microsoft's *Word* program features the Flesch Reading Ease scale, which measures the average length of words and sentences to calculate the number of years of education needed to comprehend a document. But such tools fail on several counts. For one, a long sentence that walks the reader down a path to its conclusion can be easier to follow than a muddled short sentence. And common words can be relatively long — technological or professor, for example — whereas many technical terms are short, such as meson, genome or glycan.

The common touch

Language experts generally agree that a better measure of accessibility is whether a piece of writing contains words in common usage — those that are at the front of the reader's mind, rather than tucked away in the recesses of memory. As a general principle, the greater the percentage of common words an article contains, the easier it is to comprehend.

Donald Hayes, an emeritus professor of sociology at Cornell University in Ithaca, New York, has used this principle for more than 20 years to analyse texts. He calls it lexi-

cal difficulty, and has developed a numerical scale, known as LEX, to quantify it. The scale is based on the *American Heritage Word Frequency Book*², which ranks 87,000 words by their frequency of use in textbooks, novels, magazines and encyclopaedias from US grammar schools in 1969.

Although the ranking is over 30 years old, it remains the primary word-frequency reference. 'The' is the most common word, with 'whooping' in 10,000th place. Among the scientific terms common enough to be included are 'bacteria' at 3,546, near 'pump' and 'fool'; and 'neuron', which ranks 23,595 — about as common as 'diddle'.

When calculating LEX scores, Hayes ignores the first 75 most common words as these contain little useful information. He then plots the 'cumulative proportion' of each word against the log of its rank. The cumulative proportion of, say, the 100th most common word — 'know' — is the percentage of the text made up of the words that lie between 75 and 100 in the frequency ranking. The graph (right) shows that the 1,000 most common words make up about 70% of all the words used by mothers when speaking to their children (orange line). In contrast, the same words make up only 20% of those used in the average *Nature* research article (blue line).

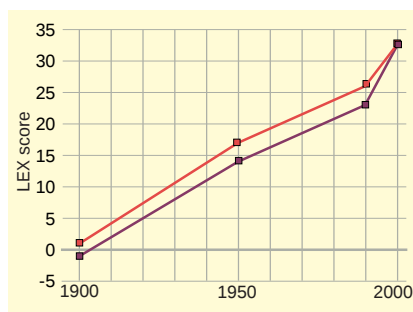
LEX values are generated by comparing the text's curve with the benchmark curve for newspapers, which have a LEX score of zero. The area under the text's curve is subtracted from the area under the newspaper curve to give the LEX value. Texts that use common words more frequently than newspapers have curves that rise rapidly, giving them a large area and a negative value; those that are skewed towards rare words end up with a positive value.

In a 1992 analysis³, Hayes found that fiction for nine-year-olds scored about -32, and a transcript of farm workers talking to dairy cows — "Let's go. Over here. You dummy, over here." — had a value of -59. Scientific papers in *Nature* and *Science* scored about 30. When *Nature* asked Hayes to repeat his analysis last year, papers in both magazines had risen to the mid-30s. This trend is not new: in the early 1900s, papers in *Science* and *Nature* had accessibility scores of close to zero (see graph, above right), similar to those of newspapers such as *The Daily Telegraph* and *The New York Times*.

Alphabet soup

What happened, says *Science*'s editor-in-chief Donald Kennedy, is that sometime after the Second World War, the number of people active in science increased dramatically, creating new subdisciplines. As they entered ever more specialized fields, new vocabularies arose. The subdisciplines of biology are among the worst for jargon. In the past 20 years, immunologists have uncovered a new world of proteins and processes, each requiring a new name or acronym. Cell-signalling research is also packed with unfamiliar terms. The average paper in *Cell*, for example, has a LEX score of about 40.

The physical sciences do a bit better. Earth scientists often use relatively common words to describe what they study, such as 'ice sheet' or 'volcano'. Specialized vocabulary exists, but there is less of it. And according to a recent unpublished study by Hayes, average papers in *Physical Review D* and the astronomy



Sign of the times: the LEX scores for *Nature* (red) and *Science* have risen steeply since 1900.

journal *Icarus* have LEX values of about 22.

The effects of an increasingly opaque literature are easy to imagine, if difficult to quantify. If opening paragraphs or abstracts are difficult to understand, researchers may miss opportunities for collaboration between disciplines. If whole papers are unclear, students get diverted to other interests and the public's fear and mistrust of science, which in part arises from difficulties in understanding new research, may increase.

Some journals are taking small steps to tackle the problem. Earlier this year, *Science* began adding one-line explanations of its papers to its table of contents. *Development* and the *Journal of Cell Science*, together with other journals published by the Company of Biologists in Cambridge, UK, have added a section to highlight half-a-dozen papers in each issue in language that is accessible to all biologists. *Nature* and *Science* have similar sections, which are complemented by longer pieces written by other academics discussing the newly published papers.

The Internet is also playing an important part in the solution. Each week, one of *Science*'s 'Perspectives' — a commentary on a published paper — in its online edition appears with links from technical terms in the text to web glossaries or sites with further information, a practice also followed by the review journals of the Nature Publishing Group. Articles in the forthcoming online journals of the Public Library of Science, a San Francisco-based organization that promotes free access to scientific literature, will be paired with lay-language summaries. And Cell Press journals now include general-interest summaries of articles in tables of contents sent out by e-mail.

But these are not perfect solutions. Scientists can be suspicious of lay summaries, fearing that they are oversimplified or inaccurate. And the Internet could exacerbate, instead of lessen, the balkanization within science. Scientists reading online are less likely to scan eye-catching figures or cogent abstracts that might entice them out of their field. And although summaries and web links lower the barrier to understanding caused by jargon and acronyms, they can't eliminate it. "These are Band-Aids," says Kennedy. "It will be

tremendously difficult to solve the problem."

It is also easy to forget why jargon is there in the first place. Technical terms are problematic for outsiders, but they are indispensable for specialists. They allow accurate shorthand for substances and processes that would take paragraphs to define. Apart from adding brief notes of explanation where space permits, the editors of top journals contacted for this article all agreed that there is little that can be done about jargon in research papers — it is here to stay.

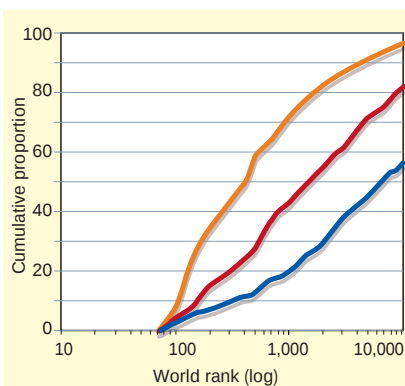
Jargon busters

But there are other ways to improve readability. "Jargon is less pernicious if you can understand what is going on," says writing instructor Judith Swan of Princeton University in New Jersey, a former biochemist who now runs workshops to help scientists to improve their papers. Swan's courses stem from a collaboration with George Gopen, a lecturer in English at Duke University in Durham, North Carolina, in which the pair developed principles of clear scientific writing by analysing published papers⁴.

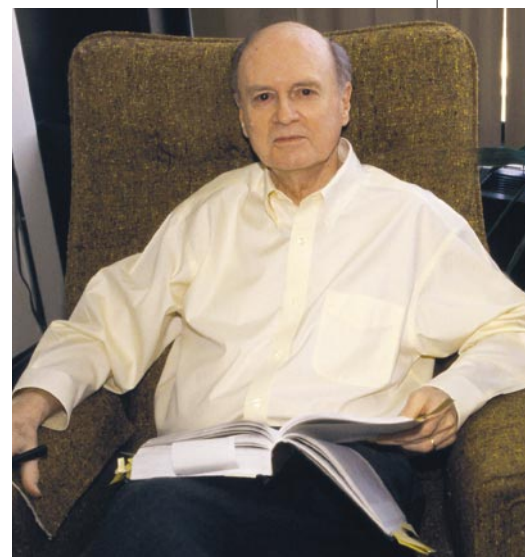
Researchers who attend the workshops expect to be told never to use the passive voice or split an infinitive, but Swan takes a different approach. "The one rule I subscribe to is that there are no rules," she says. "One doesn't follow rules, one exercises judgement."

Take passive voice. Active sentences do pack more punch, says Swan, but passive ones are sometimes clearer. For example, there is no need to begin every sentence with "We". Scientific papers tell stories about experiments and data, not scientists. Rather than use "We found the value to be x", it is fine to say "The value was found to be x", suggests Swan. "The passive is a marvellous way to hide agency when agency is not important," she says.

In general, Swan and Gopen recommend



Typical LEX curves. The red line represents a newspaper, orange is a mother talking to her baby, blue is an average research paper in *Nature*.



Brought to book: Donald Hayes has devised a scale for measuring the accessibility of an article.

focusing attention on the expectations of the readers. Linguists know that information is easier to interpret if it is placed where readers expect it to be. So, for example, when a subject is introduced in a sentence, readers expect to find a verb soon after it. Everything that comes between the subject and the verb gets little attention.

A place for everything

Take this sentence from a paper in a recent issue of *Science*: “The emergence of virulent *Plasmodium falciparum* in Africa within the past 6,000 years as a result of a cascade of changes in human behaviour and mosquito transmission has recently been hypothesized.” After the subject — “emergence” — the reader must wade through 25 words before reaching the verb — “has been hypothesized”. Readers will focus too much attention on the anticipated verb to notice the importance of the intervening material.

Swan suggests the following rewrite: “According to a recent hypothesis, virulent *Plasmodium falciparum* emerged in Africa within the past 6,000 years as a result of a cascade of changes in human behaviour and mosquito transmission.” Not only are the subject and verb snugly together — “*Plasmodium falciparum* emerged” — but now the important information occupies a key position in the sentence: the end.

The last part of a sentence is what linguists call the stress position. Readers naturally emphasize the information at the end of a unit of discourse, such as a sentence or paragraph, making it the logical spot for new information. Old information does better near the beginning of a sentence, where it grounds the reader in preparation for the mental leap to come. And the more closely the structure matches the reader's expectations, the more likely the reader is to comprehend what the author is trying to say.

Another mistake often occurs right at the start of the sentence, in the ‘topic position’. Readers expect to find some sort of bridge between sentences here. If a completely new word or phrase occupies this spot, Swan says, the reader is momentarily confused. For example, a recent paper in *Cell* begins: “We demonstrate that the tendons associated with the axial skeleton derive from a heretofore unappreciated, fourth compartment of the somites. Scleraxis (Scx), a bHLH transcription factor, marks this somitic tendon progenitor population at its inception, and is continuously expressed through

Word up: Polly Matzinger wants to see prizes awarded for the best-written scientific papers.

differentiation into the mature tendons.”

The authors unintentionally trip the reader by starting the second sentence with a brand new term — the *Scleraxis* gene. Although readers can manage this jump, it forces them to divert some of their attention from the science to the writing, particularly if the pattern recurs. To avoid this, Swan suggests sliding one of the familiar items from later in the sentence to the front to prepare the readers the new information. “This somitic tendon progenitor population is marked at its inception by the gene *Scleraxis* (Scx), a bHLH transcription factor.” Although passive, the revised sentence smoothes the prose, so readers can focus more intently on the scientific content.

Swan's advice is distributed in a variety of ways. Some institutions, such as JILA, a physics laboratory at the University of Colorado, Boulder, have offices that are dedicated to editing papers and helping scientists with their writing, and which draw on materials developed by Swan and Gopen. Swan also gives about eight scientific-writing workshops a year in the United States. The Earth sciences division of the Lawrence Berkeley National Laboratory in California hosted one workshop last year. Divisional director Bo Bodvarsson says such training is essential to a successful scientific career, particularly for students whose first language is not English, because clear communication opens doors.

Do the write thing

Despite such enthusiasm, most scientists receive no such training. Part of the reason may be that a significant minority of researchers believe that good writing cannot be taught. Among them is Christopher Miller, a biochemist at Brandeis University in Waltham, Massachusetts, known among editors for submitting clearly written papers and reviews. He says that he doesn't follow a set of writing rules, but writes instinctively and only when he is in a “writing mood”.

This instinct isn't something that Miller feels he can convey to his students, so he takes a different approach. Anyone in his lab has two chances to write a paper. “You give me a draft and it will stink, I will write a few things on it and you get another chance,” he says. If

that advice doesn't result in an acceptable paper, Miller writes it himself.

Some students and postdocs aren't interested, but those who are often produce good second drafts nearly ready for submission, he says. “It has turned out to be productive and fun.”

For those who do not have access to advisers such as Miller or Swan, professional manuscript editing services can help. Brian Leonard



Pens down: Christopher Miller believes that the instinct for good writing cannot be taught.

runs Exact Science Communications, an editing service based in Surprise, Arizona. He says that most of his clients are non-native speakers of English — others want to polish their manuscript to improve the chances of publication. In some cases the suggestion to seek professional help comes from journal editors or referees, he says.

Those editors could also do a lot more to draw better writing from contributors, says Polly Matzinger, an immunologist at the US National Institute of Allergy and Infectious Diseases in Bethesda, Maryland, and a scientific adviser to the Council for the Advancement of Science Writing, which aims to improve the quality of science journalism. Many editors already spend a great deal of time whipping manuscripts into shape, with particular emphasis on the abstract and first paragraph. But Matzinger thinks that journals should push harder and expect good writing in all submissions, possibly even rejecting papers on that basis alone. “Play it up, talk about it, insist on it,” she says. One idea would be for journals to announce an annual award for the best-written paper of the year.

With a lack of big ideas for addressing the jargon problem, bit-part solutions, such as prizes and the range of other ‘Band-Aid’ measures, are currently the best hope for promoting accessibility. Together with the techniques promoted by Swan, they should help to attract scientists to new kinds of abstracts, and keep them hooked until the end of the article. Thanks for sticking with this one. It has a LEX score of –4.1, so you don't have to be smart enough to read *The New York Times* to understand it — but it's probably too complex for cows. ■

Jonathan Knight writes for *Nature* from San Francisco.

1. Crick F. *The Astonishing Hypothesis* (Scribner, New York, 1994).
2. Carroll, J. B., Davies, P. & Richman, B. *American Heritage Word Frequency Book* (Houghton-Mifflin, Boston, 1971).
3. Hayes, D. P. *Nature* **356**, 739–740 (1992).
4. Gopen, G. D. & Swan, J. A. *Am. Sci.* **78**, 550–558 (1990).

Under new management

After decades of rule by first the military and then parties associated with the privileged élite, Brazil now has a left-wing president. David Adam considers the implications for the country's science.

If few Brazilian scientists were aware of Roberto Amaral when he took up office as the country's minister of science and technology on 2 January, most knew his name a week later. In his maiden media interview, Amaral seemed to suggest that Brazil was about to resurrect its nuclear-weapons programme — a gaffe that forced officials to issue a hasty statement ruling out such a volatile move.

It was an explosive start that played up to the worst fears of some researchers about the new government. Amaral, a political scientist and journalist who is vice-president of the Brazilian Socialist Party, was given control over research in the horse-trading that followed Brazil's election last year of its first left-wing government in four decades. President Luiz Inácio Lula da Silva, known popularly as Lula, was elected to bring about a social revolution, and the country's scientists were unsure about which way the coming winds of change would blow them.

Scientists weren't the only ones to express concern — economists and business leaders were also worried about what the new government had in store. But some six months after the new administration took over in January, there is a widespread mood of cautious optimism. Lula's government has won plaudits for its generally prudent management of Brazil's shaky economy. Scientists, meanwhile, have been wooed with a pledge to double federal spending on research within his four-year term, and to boost the number of young researchers being trained.

Controversially, however, Lula's government also wants to widen the geographical spread of Brazil's research base, currently dominated by the southern states of Rio de Janeiro and São Paulo — a move that is bound to create tensions between the current haves and have-nots of Brazilian science. The administration also faces a crunch over genetically modified (GM) crops: a court of appeal is expected to rule soon on whether the previous government's decision to license the commercial planting of GM soya is illegal, as consumer and green activists successfully argued in a lower court.

Brazil's scientists have learned to treat announcements about future riches with scepticism — too many government pledges have gone unfulfilled in the past. But for now



Men with a mission: since taking the reins as Brazil's president at the beginning of this year, Luiz Inácio Lula da Silva (left) has pledged to redistribute science funding across the country. Roberto Amaral (above) is charged with making his plan a reality.

E. PERES/AP; O. MAGALHAES/AE (INSET)

at least, the signs are promising. The US\$500-million annual federal research budget was largely spared from a series of stringent cuts, intended to help reduce Brazil's internal and international debts, announced shortly after Lula took office. "The new government has a very good dialogue with the scientific community," says Glaci Zancan, a biologist at the Federal University of Paraná and president of the Brazilian Society for the Advancement of Science.

Fruits of labour

Brazilian science has already proved that it can compete on the international stage, with the publication in 2000 of the complete genome sequence of the citrus pathogen *Xylella fastidiosa* (A. J. G. Simpson *et al. Nature* **406**, 151–157; 2000) by a

consortium in the state of São Paulo. But despite such pockets of excellence, the quality of science in this enormous country remains mixed. In terms of the number of papers published in international journals, Brazil ranks 17th in the world, contributing just 1.3% of the total. "That level of performance is, quite frankly, inadequate," says Amaral. "But our position in applied science and technology is even worse." Based on patents filed, Brazil is a woeful 43rd in the world. "While South Korea registered 3,472 patents in the United States in 2002, we only managed 113," laments Amaral.

Brazil's leading researchers agree that their country's science will be most competitive if it focuses on issues that are of key domestic importance. With more than half of Brazil's territory taken up by the

rainforests of the Amazon — a critical resource for both the country and the world — it is not surprising that the remote sensing of that region is a top priority. A new project called SIVAM — System for the Vigilance of the Amazon — became operational last summer. Combining data from satellites with sensors on aircraft and the ground, the system is the most ambitious of its kind in the world, and aims to promote sustainable development by providing real-time monitoring of issues such as deforestation, pollution and the spread of disease.

The *Xylella* genome project had similar local significance — the state of São Paulo being responsible for more than a third of the world's orange crop. "Scientists in Brazil must compete in different markets," says Fernando Reinach, a researcher at the University of São Paulo's Institute of Chemistry, who helped to coordinate the *Xylella* project.

The success of Reinach and his colleagues has spawned a series of similar projects targeting the genomes of other important agricultural pests (see *Nature* **407**, 440–441; 2000), but this investment in genome sequencing has not pleased everyone. "There is a chronic shortage of funds for science in Brazil and lots of people are very dissatisfied with the amount of money invested in genome projects," says Antônio Carlos Campos de Carvalho, an immunologist at the Federal University of Rio de Janeiro, who heads a group that aims to develop therapies using adult stem cells.

Compared with his counterparts in São Paulo, Campos de Carvalho has good reason to be concerned. His state government of Rio

de Janeiro is having a torrid time financially, and is barely able to pay university salaries. "It's crazy here," he says. "Sometimes there is money and sometimes there isn't."

Such problems are common because, as well as receiving money from the national government through several routes, Brazilian scientists also rely on their state governments for funding. Each state has its own funding agency, in name at least, but only the one operated by the state of São Paulo, called FAPESP, is truly effective. This is because under the state's law, 1% of all tax revenue goes to the agency. Other states have passed similar laws — Rio de Janeiro even pledged to double its local rival's contributions — but they are rarely enforced.

"The worst thing in Brazilian science is the instability of funding," Campos de Carvalho says. "Even if you get a grant approved, you have no idea if you will ever see the money." Still, compared with other regions, the states of Rio de Janeiro and São Paulo are lucky. Brazil is very much a country of rich and poor — and this is as true in science as in any other aspect of society, with spending on science being heavily concentrated in the two southern industrialized states.

Regional rewards

It's in trying to redistribute this wealth that Lula's government may evoke the wrath of Brazil's scientific elite. "We intend to promote a policy of social inclusiveness and break up the concentration of investment in science and technology," Amaral says. "We aim to create incentives to help researchers establish themselves in their own geographical regions." Portions of the federal budget may be reserved for research groups outside Rio de Janeiro and São Paulo.

The proposal has split the research community. Zancan, for one, supports it: "It is a mistake to leave emerging groups to their own resources," she says. Unsurprisingly, successful groups in São Paulo and Rio de

Janeiro argue that the move will undermine Brazil's efforts to become internationally competitive in science. "Lots of money already goes to the regions and much of it is just wasted," claims Reinach. "It's a very tricky problem, but the idea goes against the culture of merit and means that good science in São Paulo will not get funded."

Another controversial issue is the stance of Lula's government on transgenic crops. In opposition, Lula was seen as an opponent of GM agriculture. But his cabinet is split on the issue, and he has now set up a commission to produce an official declaration. Last month, the government decided to allow the GM soya currently being grown in the south of country — in defiance of the court-imposed moratorium — to be sold until 31 January next year.

Despite such potential flashpoints, most scientists remain reasonably confident that Lula's government will build on its predecessor's achievements in science policy, which involved channelling more money into research. For example, an innovative new funding mechanism, under which taxes from various industries including oil and information technology are fed back into related research, currently provides about US\$400 million a year, over and above the federal research budget. "One of the important things that has happened over the past few years is that Brazil has built a system," says Hernan Chaimovich, who heads the chemistry department at the University of São Paulo. "The new government will shift the emphasis, but the system is set."

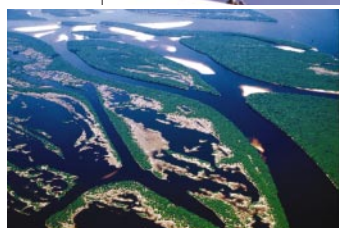
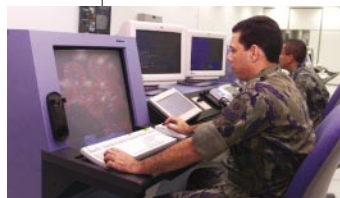
A stiffer challenge will be turning investment in basic science into innovations that benefit local communities and the economy as a whole. Brazilian industry has tended to bring in expertise rather than trying to develop it locally, so the government is now considering redirecting a portion of federal funds towards research projects with commercial potential. Examples include exploiting the biodiversity of the Amazon region and developing alternative fuels — Brazil already has decades of experience in running cars on ethanol derived from sugar cane. "We want to develop the technical and scientific tools to contribute to long-term national development," says Amaral.

The ultimate test for Lula's government, however, will be whether it can improve the lot of the poor citizens whose votes propelled it into office. And when the next elections come round, some scientists fear that investment on science may lose out to shorter-term priorities. "No kids in the United Kingdom had to starve to let John Sulston sequence the human genome, but in Brazil we still have to worry about feeding people," says Reinach. "Science here always runs the risk of appearing to be a luxury."

David Adam, until recently a news and features writer for *Nature*, now writes for *The Guardian* in London.

L. E. PEREZ/CCSIVAM

Brazil's SIVAM project to monitor the Amazon forest combines data from ground, airborne and satellite sensors.



Embryos aren't essential to stem-cell research

Sir — Your Editorial “Disease insights from stem cells” (*Nature* 422, 787; 2003) gives the impression that most US scientists support human embryonic stem-cell research. Not everybody would agree; I would like to state a less reported view.

The suggestion made by some that cloned human embryos will provide otherwise unobtainable cells for disease research misleads the public. There are available countless cell lines derived from a wide variety of normal and diseased human tissues. A greater concern is the practice of understating the scientific challenges of using human embryos to produce mature adult tissues *in vitro* or *in vivo*. For embryonic stem cells to make adult tissues, they must first be converted into adult stem cells. All the fuss over embryonic stem cells has damaged public enthusiasm for any type of stem-cell research.

The US Congress should target as much as \$200 million of funding for research over the next five years to realize the promise of stem cells for human therapies. This would provide enough money for several projects at each of the 18 National Institutes of Health (NIH). This cost is a small fraction — less than 1% — of the present NIH budget (more than \$27 billion). Both adult stem-cell research and non-human embryonic stem-cell research should be supported. Non-human research might one day yield methods of reprogramming human cells to early developmental states without creating or destroying human embryos.

Congress should motivate stem-cell scientists to do what is in the best interest of the public. This is a necessary tension to ensure the long-term quality of the publicly funded research enterprise. As for human therapeutic cloning, the US public has been waiting a long time for leadership that is both moral and based on good science. In the long run, the best interest of the public will also be the best interest of scientists.

James L. Sherley

Biological Engineering Division, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

US federal funding ban puts babies at risk

Sir — We wholeheartedly agree with your Editorial “Reining in assisted reproduction” (*Nature* 422, 647; 2003) that more research is needed to identify and evaluate the risks that assisted reproduction may pose to children and their mothers.

Politics hindering SARS work

Taiwan has been left to fight its outbreak with little help.

Sir — Your News story “Taiwan left isolated in fight against SARS” (*Nature* 422, 652; 2003) highlights a problem scientists here have faced for years. I was invited to a workshop on mathematical ecology in Trieste in 1988, arriving only to find myself excluded from the list of participants and not allowed to give a talk.

Upon inquiry, I was told by the local organizer that “officially” I was not present at the workshop, co-sponsored by the International Atomic Energy Agency and UNESCO, owing to a protest by the Chinese Embassy in Rome after I had been invited. Apparently, unlike scientists from anywhere else, all Chinese scientists had to be recommended by the Chinese government to attend that workshop, and since Taiwan was considered to be part of China, I could not attend without permission from the Chinese government.

The SARS epidemic illustrates that being politically correct in scientific matters does more than just inconvenience a few Taiwanese scientists. Indeed, with the official global death toll from SARS at 643 by 19 May and rising fast, one can only speculate how many lives could have been

saved had the health officials in Hong Kong been a little less worried about being politically correct in dealing with the emergence of this epidemic.

It is incredible that, even now, with the SARS death toll in Hong Kong having reached 247 by 18 May, officials there still refuse to link the problem with Hong Kong's proximity to Guangdong Province in China, where similar symptoms have been appearing since November. Almost all early infections in Hong Kong can be traced to recent travellers from China or people in close contact with them.

The World Health Organization sent two epidemiologists to Taiwan to assess the situation on 3 May. Meanwhile, left to fight SARS alone, Taiwan had a threefold rise in the number of cases in April and by 19 May had had 344 probable cases and 40 deaths.

Unfortunately, the Hong Kong government seems more preoccupied with how it looks to China and the rest of the world than with saving lives.

Ying-Hen Hsieh

Department of Applied Mathematics, National Chung Hsing University, 250 Kuo-Kuang Road, Taichung, Taiwan 402

In the United States, a ban on federal funding for embryo research makes this work especially difficult. Lifting this ban would both enable more outcome studies and ensure that research on new treatments would be subject to federal supervision and would be conducted in accordance with human-research subject protections under the common rule.

In the United Kingdom and other European countries, patients benefit from state-supported healthcare systems that subsidize and regulate the treatment for infertility and assisted-reproductive technologies. Unfortunately, most patients in the United States pay for their expensive treatments out of their own pockets. Sometimes, parents' desire to maximize their chances of having a baby and their interest in minimizing risk to themselves and their offspring are in conflict.

We have doubts as to whether a system analogous to the United Kingdom's Human Fertilisation and Embryology Authority is appropriate in the United States, where citizens fervently guard their rights to privacy and reproductive choice.

Many would see a government body with the power to dictate treatment choices as an unwarranted imposition and intrusion, especially because the

government does not subsidize payment for their treatment.

Proactive, responsible self-regulation is the key to ensuring patient safety. The American Society for Reproductive Medicine (ASRM) guides medical practice and research and keeps members abreast of the ethical issues arising daily in reproductive medicine; the Society for Assisted Reproductive Technology (SART) sets the standards for assisted-reproductive technology clinics, including laboratory certification, reporting of clinic success rates and compliance with the ASRM guidelines. This voluntary yet effective procedure reflects the dedication of our members to assuring safety, efficacy and confidentiality for all our patients.

We hope that the US government will one day fund research in this field, with accompanying protection. In the meantime, we appreciate our international colleagues whose research informs our own work.

Sandra Carson*, Robert Brzyski†

**President, American Society for Reproductive Medicine (ASRM), 1209 Montgomery Highway, Birmingham, Alabama 35216, USA*

†President, Society for Assisted Reproductive Technology (SART), 1209 Montgomery Highway, Birmingham, Alabama 35216, USA

Making your mind up

Neurons with sustained activity could help us to understand cognition.

Cortex and Mind: Unifying Cognition

by Joaquín M. Fuster

Oxford University Press: 2002. 314 pp.

\$49.95, £47.95

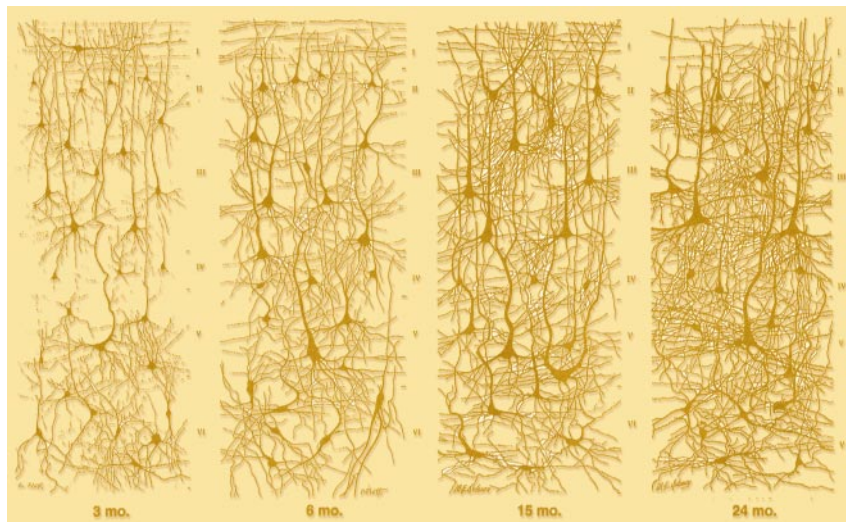
Kevan Martin

"Human cognition may like the winged horse take at times its flights toward the stars and forget Earth." This was Charles Sherrington's own flight of fancy as he pondered the ancient question of the relationship between the working of the brain and the working of the mind in his Gifford lectures of 1937–38. How close are we now to an answer? Joaquín Fuster provides a progress report in *Cortex and Mind*, in which he considers several aspects of cognition: perception, memory, intelligence, language and attention. What is striking from his account is how much we now rely on data from experimental animals, particularly primates, to understand the physical basis of human cognition — more even than Sherrington, who was perhaps the greatest of all behavioural physiologists.

Those familiar with Fuster's previous book *Memory in the Cerebral Cortex* (MIT Press, 1994) will know that he discovered in the prefrontal cortex of monkeys a group of neurons that sustain their activity between a sensory cue and a delayed motor response. The existence of such 'working memory' cells had long been predicted, and soon afterwards Hiroaki Niki found a second class, which accelerate their firing over the delay period.

These two classes are intermixed in most of the prefrontal cortex and are also found in parietal cortex. It is not yet clear whether such sustained activity arises through reverberation in the recurrent cortical circuits, as proposed by Lorente de Nó and currently the theorist's favourite, or through mnemonic capabilities at the level of single neurons, as suggested by *in vitro* experiments by Alexei Egorov and colleagues, and no definitive answer is expected soon. But the question is nevertheless important: how do our cognitive processes map onto the brain's circuits?

For Fuster, this question is decided by his view that the brain is a neural network, currently best represented by connectionist models involving parallel distributed processing. The essential currency of cognition for Fuster is the 'cognit', a word he coins here to describe an item of knowledge that is stored in a network. Each node of the cortical neural network stores yet more elementary representations that make up a cognit. Not all cognits are equal. Symbols, for example, are presented in high-ranking cognits, themselves produced by the convergence of



Growth factor: neurons develop rapidly in the human cortex between 3 and 24 months of age.

cognits from lower levels. In other words, the essential transactions underlying cognition involve the activation of different networks, arranged in hierarchies, that represent the building blocks of cognition. So, unlike Sherrington with his 'million-fold democracy' of neurons, Fuster proposes an isomorphism between the processes and the structures of mind and of the cerebral cortex.

This is a nice idea, but how well does it fit the facts? Fuster's chosen model of the parallel-distributed-processing network has been widely used, especially by cognitive psychologists. Such artificial neural networks have multiple layers, learn associatively, provide a complete output from incomplete input data, and show 'graceful' degradation in the face of partial damage. These are all presented as being satisfyingly brain-like.

But in truth, the performance of artificial neural networks is fragile in real-world environments. They seem ill-suited to representing Fuster's cognits, because the nodes of these networks usually represent functions with continuous values and their outputs combine with those of all other nodes to compute the output of the network. The nodes do not represent discrete bits that can simply be composed into well-defined higher-order constructs. Indeed, artificial neuronal networks can approximate any arbitrary non-linear function without resorting to extensive hierarchies, yet hierarchies seem to be a fundamental organizing principle in brains. Artificial neural networks also scale notoriously badly, so most successful simulations have usually used networks with substantially fewer than 1,000 'neurons', in contrast to the 100,000 neurons contained in each cubic

millimetre of neocortex. How we learn effectively and stably with such large networks is a mystery. Most models propose simple feed-forward networks, which do not capture the essential recurrence and local processing of the neocortical circuits, nor do they capture the temporal aspect of processing, which is a feature of perception and action.

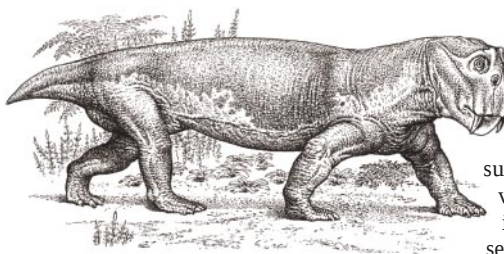
The capacity of the brain to excite itself is a fundamental necessity for cognition. And one of the key capacities that working memory provides is the possibility to delay a choice of action until the various options have been weighed up. This 'executive' model of prefrontal cortex has been popular since Phineas Gage survived a horrific mining accident in 1848 to become the first clinical exemplar of the consequences of damage to the prefrontal cortex. But as we 'ponder', 'deliberate' or 'examine' (all words deriving from Latin roots meaning 'to weigh'), can we find an analogous 'weighing' in the activity of neurons in our prefrontal cortex, or are the brain computations involved in these cognitive processes carried out by symbolic manipulations reminiscent of our use of numbers and letters? Exactly how do neurons make up one's mind?

Horace Barlow once proposed that the rate of neuronal activity reflects the probability of a certain stimulus being present. Decision-making requires that these probabilities be weighed, and one intriguing recent suggestion, by Michael N. Shadlen, is that this occurs in a second stage of processing in which sensory responses are accumulated from pools of neurons over time to form a single quantity, for example the logarithm of the likelihood ratio. The neurons that best

fulfil this role are none other than those discovered by Fuster and Niki. If their persistent activity in the absence of a sensory cue is indeed the step of calculating a single decision variable based on information from several sources, then neurophysiologists have actually watched neurons making up the monkey's mind. What determines the moment of decision is not yet known, but just as 'decide' once meant to cut off, or bring to an end, so these neurons do indeed stop their activity when the decision is made.

There is a strong argument that we have made such great progress in understanding the neural basis of cognition only because neurons, and the networks that they form, compute in an analogue style. We can get an idea of the underlying computations by measuring the activity of single neurons, or the strength of the functional magnetic resonance imaging signal. It seems fantastic, but Fuster's progress report dares us to believe that the patterns woven by Sherrington's "enchanted loom", the cerebral cortex, are now well on the way to being understood. ■

Kevan Martin is at the Institute of Neuroinformatics, University of Zurich/ETH, Winterthurerstrasse 190, 8057 Zurich, Switzerland.



Exit stage right — even though *Lystrosaurus* survived the extinction at the end of the Permian.

us how in the 1990s the evidence began to emerge that the species replacements marking the Permian–Triassic transition were also sudden, and hence were probably caused by some environmental trauma. He is describing both a geologically sudden event and a rapid transformation in our ideas about the Earth's past.

As a result, the book is partly historical in nature. It describes how the British geologist R. I. Murchison (himself a catastrophist) defined the Permian rocks of Russia in about 1840, and how Lyell and Darwin challenged the idea of mass extinctions by arguing that apparently sudden transitions in the fossil record were the result of gaps in the evidence, which created illusory jumps between one system of rocks and the next.

The triumph of darwinism ensured that catastrophist explanations were marginalized until they were revived by the asteroid-impact theory for the end of the Cretaceous. Even then, many palaeontologists resisted, arguing that the dinosaurs were declining anyway, so the impact only finished a job that had already been started by gradual environmental changes. At the time, knowledge of the Permian–Triassic transition was so limited that gradualism still seemed plausible here, too. Benton provides a graphic account of how more recent evidence has piled up, including his own experiences fossil hunting in Russia, making a catastrophic explanation inescapable.

There is one important twist in the story, however: Benton finds little support for the possibility that the Permian extinction was caused by an extraterrestrial agent. Wild theories about periodic bombardments by asteroids have not stood the test of time: the Permian event was probably triggered by massive volcanism, which injected poisonous gases into the atmosphere, both directly and by triggering the release of methane from deep-sea hydrates. Some geologists think that volcanism also played a role at the end of the Cretaceous. Significantly, Benton concludes by considering the implications of the latest, man-made mass extinction, asking what light the earlier events can throw on the potential for survival of modern species.

The historical aspect of Benton's book raises some intriguing questions. Many early catastrophists postulated the involvement of extraterrestrial agents — a comet was sometimes invoked as the cause of Noah's flood.

But such ideas went out of fashion in the mid-nineteenth century, and later catastrophists, including Murchison, favoured explanations based on the supposedly more intense geological activity in the young Earth. The asteroid-impact theory of dinosaur extinctions seems to parallel some of the earliest speculations, but Benton has redressed the balance by favouring internal causes.

My one criticism of his account is that he accepts too readily the assumption that Lyell and Darwin marginalized all support for discontinuity in the Earth's history. There were few outright catastrophists left by around 1900, but many still believed that the history of life had been punctuated by environmental transitions far more rapid than anything observed in the recent past.

The real triumph of gradualism came with the modern darwinian synthesis of the mid-twentieth century, and even then it was confined to the English-speaking world. Benton notes that British and US palaeontologists of the 1950s ignored the catastrophism of Otto Schindewolf. But we need to recognize that German palaeontologists such as Schindewolf were continuing a long-standing tradition that had proved far more robust than our modern, Darwin-centred histories acknowledge. The fact that modern catastrophists do not see a link back to that tradition tells us about the effectiveness of the neo-lyellian interlude of the mid-twentieth century. ■

Peter J. Bowler is in the Department of Social Anthropology, Queen's University Belfast, Belfast BT7 1NN, UK.

Suffocated or shot?

When Life Nearly Died: The Greatest Mass Extinction of All Time

by Michael Benton

Thames and Hudson: 2003. 336pp.

£16.95, \$29.95

Peter J. Bowler

Whatever hit the Earth at the end of the Permian period certainly struck hard, killing 90% of living species. Compared with this, the extinction at the end of the Cretaceous period was comparatively minor, with only a 50% death rate. Yet the latter event is much better known, because among that 50% were the last of the dinosaurs. Partly for this reason, Michael Benton uses the event at the end of the Cretaceous as an introduction to his account of the Permian extinction — he wants us to realize how limited it was in comparison with what he intends to describe.

But there is a deeper reason for linking the two episodes: Benton wants to show us how the catastrophist perspective has re-emerged in modern geology and palaeontology. He argues that the theory of catastrophic mass extinctions was widely accepted in the early nineteenth century, but was then driven underground by the gradualist perspective of Charles Lyell's uniformitarian geology and Darwin's theory of evolution. Only in the 1970s was catastrophism revived, through the claim that the dinosaurs were wiped out when an asteroid hit the Earth. Benton shows

Hooke, life and thinker

London's Leonardo: The Life and Work of Robert Hooke

by Jim Bennett, Michael Cooper, Michael Hunter & Lisa Jardine

Oxford University Press: 2003. 240 pp.

£20, \$35

David R. Oldroyd

Some devotees of Robert Hooke have regarded him as Britain's greatest scientific genius of the seventeenth century, the range of his interests and achievements being hard to conceive. He is a fruitful subject for historical enquiry as he left behind him a large archival trail, and, with his polymathic interests, he has attracted much attention. A good general overview, *Robert Hooke* by Margaret 'Espinasse' (Heinemann), was published in 1956. Since then, studies of Hooke have expanded greatly to the point where we have a detailed knowledge of the man, although not all within the pages of a

single volume. *London's Leonardo* contains four highly competent and complementary essays, which go a long way towards providing a definitive account of Hooke, while leaving open the road (or preparing the way) for a full intellectual biography.

Hooke was wealthy at his death, much of his money having come from his work helping to resurvey London after the Great Fire of 1666. In his essay, Michael Cooper describes this work pleasantly and informatively. That Hooke should have embarked on it when he was already fully occupied with his scientific work for the Royal Society is remarkable and bespeaks his devotion to London and its inhabitants. There were many problems. With street widening, residents had to be compensated fairly for the land they were to lose. Buildings had different owners on different floors, and some structures had 'interleaved' with their neighbours. An accurate survey was needed, and it relied on instruments, some devised by Hooke, that were an integral part of the 'scientific revolution'. Hooke's contributions to the survey were substantial.

Jim Bennett's fine paper, which is profusely illustrated, deals with Hooke's instruments and inventions more generally, revealing their extraordinary range and ingenuity: time-pieces, air pumps, telescopes and microscopes, meteorological and oceanographic instruments, the universal joint and many other items. Hooke believed in the use of instruments to enhance the senses, as can be seen from his controversy with the Polish astronomer Johannes Hevelius, who still advocated naked-eye instruments for astronomy. Hooke was clearly on the winning side. Everyone knew that optical instruments had imperfections, and Hooke applied himself to the endless task of their improvement.

Michael Hunter writes about Hooke's philosophy of nature and his ideas on scientific method. Regarding the latter, Hooke was not a baconian inductivist (nor, indeed, was Bacon), but rather a hypothetico-deductivist. Although Hooke made some use of baconian tables of 'presence', 'absence' and 'degrees', he gave a clear example of the formulation and



Instrumental to his success: Hooke relied on optical devices such as this compound microscope.

Art Science in site



Taking issue: *Happy Hour* by Fernando Arias (left) examines AIDS treatments; Daniel Lee focused on evolution for *Cheetaman* (middle); and Annie Cattrell's *Capacity* was inspired by the breath of life.

The website scicult.com is a science-related contemporary art gallery — and an act of love. The small group of 'sci-art' specialists who launched it earlier this year are idealists, committed to promoting a quality marriage of art and science.

The group has already signed up 20 significant artists, including Annie Cattrell and Fernando Arias, some of whose work is shown here. The art is exhibited in the online gallery, and some pieces will eventually be available for sale.

But scicult.com is more than a gallery. It publishes an expanding range of intelligent features about contemporary sci-art, and has

longer-term plans to develop an 'introduction service' for scientists and artists who seek collaborating partnerships. It is also in the process of acquiring a permanent, real-world gallery in which it can exhibit more experimental works.

The website is attractive and functional. Artworks are well displayed against a dark-grey background and can be enlarged with a click of the mouse. The features are timely and well-written, but suffer the plague of many web pages designed primarily for visual impact: the text, reversed out white on dark grey, is a strain to read on the screen.

Alison Abbott

◆ www.scicult.com

testing of hypotheses in science. He proposed the idea of pole-wandering to account for cyclical interchanges of the levels of land and sea (to explain the presence of inland fossils). Such movements in the position of the geographic poles, if they occurred, would, over time, produce changes in the direction of the meridian at any given locality. Hooke then suggested astronomical methods for the accurate determination of the meridian, which should be measured over a period of years to look for changes. A first attempt at determination failed because of poor weather and the idea was not pursued, being pushed aside by Hooke's manifold other activities, but the hypothetico-deductive method was clearly enunciated.

This example, in a way, renders superfluous historians' worries about what Hooke meant by what he mysteriously called 'philosophical algebra', presumably some kind of 'routinizable' procedure for conducting science. Of course, knowing about the 'form' of scientific method tells us little about how Hooke's creative process worked. Hunter, unlike another Hooke aficionado, Steve Shapin, eschews discussion of the significance of Hooke's social status for his scientific practice. Rather, Hunter gives an excellent exposition of Hooke's *Micrographia*, which links back to the discussion of instruments, and further illustrates his procedures.

Lisa Jardine's paper is less precisely focused than the other three. She explicates details of Hooke's relations with Robert Boyle, and writes about Hooke's work on pressures, the magnitude of subterranean gravitational attraction and geology. But she is chiefly interested in his health and his self-medication (recorded in his diary), which eventually more or less killed him. Hooke left no will, and his family fell on his fortune after he died. They were not interested in preserving his name, so for many years he was a rather forgotten figure (Jardine suggests). But his time has come: the comprehensive bibliography of *London's Leonardo* shows just how many works have been written about him since 'Espinasse's biography.

This prompts a thought. People's interests can often be judged by their libraries. Hooke's printed library sale catalogue survived, and some years ago I attempted an approximate classification of his books. The number of literary items (languages, grammar, philology, poetry, plays, epigrams and biographical works) easily exceeded the number in any of the categories of mathematics, astronomy, logic, physics, architecture, machines and so on. Is there perhaps another Hooke to be explored: the man of letters?

David R. Oldroyd is in the School of History and Philosophy, University of New South Wales, Sydney 2052, Australia.

Into uncharted waters

What was your first experiment as a child?

An attempt (aged 6) to reproduce the conditions of Hell among wooden toys and rubber balls inside the hotplate compartment of the imposing tiled stove in our sitting room. The results hinted that this might be possible, although exact data could not be obtained. The thermometer had melted by the time our neighbours rescued us and the fire brigade arrived. A lesson learned was that you should consider success with all its implications at least as carefully as potential failure.

Who has been the most important mentor in your career?

My junior-school teacher in Budapest, György Ercsényi, who opened my eyes to natural sciences and taught me to appreciate the power of logic without hurting fantasy.

What makes a good scientific mentor?

Being perfectly imperfect while also broad-minded and generous. It is also useful to have something worthwhile to rebel against, and to have something beautiful to cling on to. I was very lucky with my supervisors and mentors: Vilmos Csányi (a chemist who wrote on evolution and behaviour); F. Bruno Straub (who in 1988–89 became the first non-communist president of Hungary); Péter Friedrich, who succeeded Straub at his institute in Budapest; Louise Johnson, who got me into structural biology in Oxford; and David Phillips, who taught me about structures in science.

What single scientific paper or talk changed your career path?

The most recent change was brought about by two lectures: one by Herman Winnick (Stanford Synchrotron Radiation Laboratory) and one by Björn Wiik (DESY, Hamburg), who talked about plans to build an X-ray free-electron laser within our lifetime — plans that might now come to fruition.

What literary character would you employ as a postdoc?

The Queen of the Night. Puck automatically comes in the package.

What book currently resides on your bedside table?

Uncle Petros and Goldbach's Conjecture by Apostolos Doxiadis, and Adlard Coles' *Heavy Weather Sailing*.

What music heads the playlist in your car or lab?
Swedish jazz, Finnish tango, Talking Heads and music from Bartók and Mozart. Weird, isn't it?

Assuming the dead can be raised and/or time travel exists, who from the world outside science would you most like to have dinner with?

Roger II, the Norman king of Sicily, and his adviser, Abu Abdullah Mohammed Ibn al-Sharif al-Idrisi, who between 1139 and 1154 produced an amazingly accurate picture of the known world. *The Book of Roger*, completed in 1154, starts with the following passage: "The Earth is round like a sphere, and the waters adhere to it and are maintained on it through natural equilibrium which suffers no variation." What a beautiful description of gravity. Columbus knew about the book, and rumour says that he had it on his boat.

Where and when would you most liked to have lived or worked?

Here and now. We are probably witnessing one of the biggest changes in the history of biological evolution — a genome is reading itself (and others) a bit like we read books, and then storing this information in unconventional forms for later use.

What one thing would you rescue from your burning laboratory?

Nothing. Such an occasion is an opportunity.

What do you do to relax?

Work on the boat and sail her hard.

What would you have become, if not a scientist?
Some other type of hedonist.

Do you have a burning ambition to do or learn something of no practical or immediate value? If so, what?

Yes I have, and I have been practising it every day in the lab.

Under what conditions do you have your greatest and most inspired ideas?

When teaching or when preparing to teach.

Whose graduate student would you most like to have been?

Max Perutz for scientific supervision and Erasmus of Rotterdam as moral tutor. It is good to have it thick.

What's the best piece of advice you've ever received?

Don't become a missionary. Diversity is beautiful.

What's the one thing about science that you wish the public understood better?

First: you cannot legislate against evolution. Second: look what happened to the dinosaur and remember it.



Janos Hajdu

Janos Hajdu is professor of biophysics at the Biomedical Centre in Uppsala, Sweden. He loves the North, the sea, and does not want to die before seeing a living cell in molecular detail. He may need to prolong his life indefinitely.

Is there a 'tyranny of reductionism' in how scientists are trained today?

Things are changing. There is no such thing as botany, inorganic chemistry or photon physics any more. The big picture is in a mixture, and one needs to be versed in many of the former 'disciplines' to achieve a single goal. I think it is back to the Renaissance, but it is not that easy to turn into a Leonardo (and money does not come from the Medicis or the Sforzas either).

You've just been told (in confidence) that the world will end tomorrow. What do you do next?
I could consider doing that quick experiment that has the potential to destroy the whole world if it misfires but that would save it, of course, if it did not. I could also call my friends, take out the case of vintage champagne, and watch the show together all the way.

What's the most interesting thing in your fridge?
Fermented herring and a whale steak. Don't ask me how they got there. Any takers?

What music would you have played at your funeral?

My interest in music would probably be at its lowest on the occasion.

How would you like to be remembered?
Frequently.

The retina's fancy tricks

Richard H. Masland

The vertebrate eye is far more than a passive receptor for visual information. The microcircuitry in the retina can, for instance, carry out the job of distinguishing object motion from background motion.

When the first light-sensitive amoeba drifted down a stream, it encountered one of the fundamental problems of vision, which is that the world doesn't sit still. Trees move in the breeze, grass rustles, the sun and stars drift across the sky. The most basic use of an organism's light-sensitivity is to orient itself to (or from) a light source. But how, with so much moving clutter, could the amoeba have charted its position?

Fast-forward in evolutionary time to vertebrates, and things become worse yet. The image of even a peaceful world dances in great whoops and swirls when the eye moves. And it moves whether we command it to or not: the eyes make incessant, unconscious drifts, even when we stare fixedly at a single point. Given this visual dance, how can we see anything except a blur? As described on page 401 of this issue¹, an experiment carried out by Ölvéckzy *et al.* reveals one of the ways that the visual system does it. The process involves a clever piece of image processing (Fig. 1) — remarkably, that processing occurs in the neural microcircuits of the retina.

The computational task is to distinguish actual motion of an object in the world from motion across the retina caused by the fixational eye movements. The experiment was to record, during one of two conditions, from one of the output neurons of the retina, a retinal ganglion cell (whose collectively bundled axons make up the optic nerve). In the first condition, the stimulus matched the retinal input generated by an eye movement: all of the visual input — for the purposes of this experiment a black-and-white striped grating — moved across the retina with a single trajectory. The second condition simulated an object actually moving in the world: there was motion contrast, such that a test object moved in one direction and the larger, surrounding visual input moved independently.

When Ölvéckzy *et al.* compared the two conditions, the result was very simple. If an object moved independently of its background, the neuron sent a signal down the optic nerve to the brain. If everything moved together, there was no response. In effect, the circuits of the retina figure out that uniform, undifferentiated movement is only the result of eye movement, and the retina aborts its report to the brain. This behaviour was



Figure 1 **Catching the action.** In vision, certain objects pop out of the general scene with a perceptual vividness that overshadows others, even though the competing objects may have the same colours and brightness. One of the neural mechanisms involved has been revealed by Ölvéckzy *et al.*¹. Movement of an object relative to its background creates correlated firing in a group of retinal neurons. This neural activity serves to signal the unity of the moving object to the brain, with a final result artificially illustrated in the version of the image on the right.

observed for only a subset of the known types of ganglion cell. So a more complete survey is needed, as is a rationale for why it would be good for particular types of retinal ganglion cell and not others to have the local motion-detecting behaviour. But the fundamental principle is now established.

How is this trick accomplished? One of the surprises of the past few years has been the unsuspected complexity of the retina's microcircuitry². Various kinds of neurons, such as bipolar and amacrine cells, intervene between the photosensitive cells (the rods and cones) and the ganglion cells, and they shape and compress the raw information detected by the photoreceptors for efficient transmission to the brain. The bipolar cells come in roughly a dozen types; these are the retina's through-pathways, an array of parallel channels conducting different types of information from photoreceptors to ganglion cells. Along the way, information is shaped by the even more diverse array of amacrine cells, inhibitory neurons that come in at least 29 varieties (Fig. 2, overleaf), each with a limited set of synaptic partners. Ölvéckzy *et al.* went on to record from one

type of amacrine cell, termed 'polyaxonal' because it has many axon-like processes that spread widely across the retina^{3,4}. The cells responded with a timing that would well suit them to provide the 'blanking' signal that aborts the firing of the retinal ganglion cell when the visual input moves coherently across the retina.

Where does this take us? Among other things, it encourages a search for even more sophistication in the retina's computations. Other amacrine cells provide feedback signals for retinal gain control — a crucial function that adjusts the retina's sensitivity to match the ambient illumination and contrast. A single amacrine cell, the starburst cell, appears to compute the fundamental asymmetry that enables some ganglion cells to report the direction of a moving stimulus^{5,6}. Another mechanism compensates for the ballistic eye movements that hurl the eyes from object to object⁷. There are now hints that not just movement but also the spatial pattern of the moving target affect the retina's responses⁸. The great turn-of-the-century anatomists recognized the retina as one of evolution's masterpieces. Our

BETH A. KEISER/AP

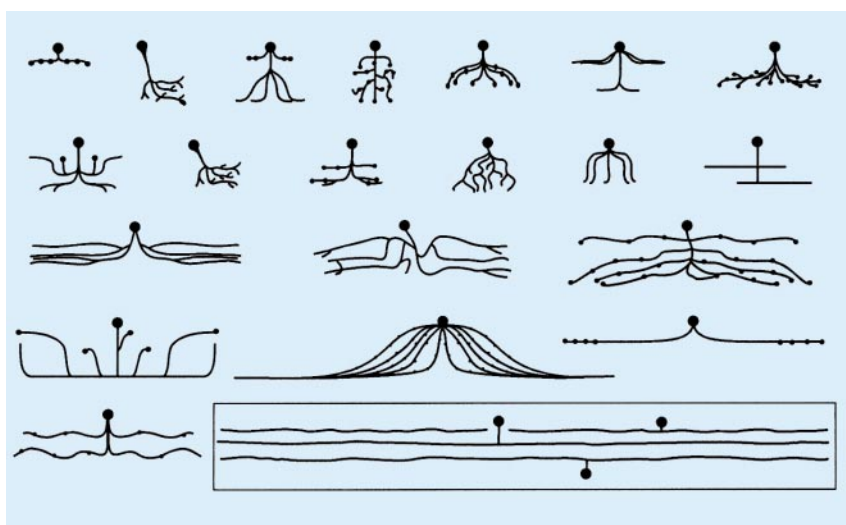


Figure 2 Information-processing machinery in the retina. As well as the input (rods and cones) and output (ganglion) cells, the retina contains a hugely diverse population of neurons. Each cell type is thought to have a specific role in vision⁹. The most extreme diversity is exhibited by amacrine cells, a collection of which is depicted here: their differing shapes and sizes are reflections of their wide variety of synaptic patterns. The widely spreading 'polyaxonal' amacrine cells, which come in several subtypes and are highlighted at lower right, are the type that Öveczky *et al.*¹ suggest provide the mechanistic basis of local motion detection. (Data for the rabbit, adapted from ref. 2.)

understanding of its signalling repertoire is finally beginning to catch up.

Richard H. Masland is at the Howard Hughes Medical Institute, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.

e-mail: masland@helix.mgh.harvard.edu

1. Öveczky, B. P., Baccus, S. A. & Meister, M. *Nature* **423**, 401–408 (2003).
2. Masland, R. H. *Nature Neurosci.* **4**, 877–886 (2001).

3. Vaney, D. I., Peichl, L. & Boycott, B. B. *Proc. R. Soc. Lond. B* **235**, 203–219 (1988).
4. Famiglietti, E. V. J. *Comp. Neurol.* **316**, 422–446 (1992).
5. Euler, T., Detwiler, P. B. & Denk, W. *Nature* **418**, 845–852 (2002).
6. Fried, S. I., Münch, T. A. & Werblin, F. S. *Nature* **420**, 411–414 (2002).
7. Roska, B. & Werblin, F. *Nature Neurosci.* doi:10.1038/nn1061 (2003).
8. Chiao, C.-C. & Masland, R. H. *Invest. Ophthalmol. Vis. Sci.* **Suppl.** 3255 (2003).
9. Masland, R. H. *Curr. Opin. Neurobiol.* **11**, 431–436 (2001).

Astronomy

New direction for γ -rays

Eli Waxman

The origin of energetic γ -ray bursts is still unknown. But the detection of polarization of the γ -rays provides fresh insight into the mechanism driving these powerful explosions.

Gamma-ray bursts (GRBs) are short flashes of γ -rays, typically tens of seconds long¹. First detected in the 1960s, GRBs are observed at a rate of roughly one per day. Although their sources are known to reside in distant galaxies, several billion light years away, what these sources are remains a mystery. But a new clue is provided by Wayne Coburn and Steven Boggs², who, on page 415 of this issue, report the detection of polarization — a particular orientation of the electric-field vector — in γ -rays from a burst. This discovery may shed light on the identity of the sources of GRBs, as well as on the mechanism by which the γ -rays are produced.

The huge energy release associated with a GRB is thought to be created by the gravita-

tional collapse of a star to form a black hole or neutron star³. The contraction causes gravitational energy to be released. The typical energy output of a GRB corresponds to the conversion of about 1% of the Sun's mass into energy: in comparison, the energy output of an atom bomb is equivalent to the conversion to energy of about 1 gram of matter.

The energy released in the collapse seems to be carried away from the source in the form of a highly relativistic jet (Fig. 1), in which particles move at nearly the speed of light. As the jet reaches a radius of about 100 million kilometres from its source, part of this energy is converted to γ -rays, which become the GRB. At a later stage still, as the jet expands to a scale of 10 billion kilometres, an 'afterglow'

of lower-energy radiation, at X-ray, optical and radio wavelengths, is produced.

The detection over the past few years of GRB afterglows⁴ provides strong support for this picture. But as afterglow radiation is produced at a large distance from the collapsing object, key questions remain unanswered³: what, for instance, is the nature of the collapsing object? The most popular candidate is a massive star, about ten times the mass of the Sun, whose life ends with the collapse of its core. How the gravitational energy released becomes a relativistic jet and how jet energy is converted to γ -rays are also not well understood.

From earlier observations of the γ -ray spectrum of GRBs¹, it was concluded that the most likely mechanism for γ -ray production is 'synchrotron emission' — the emission of radiation by highly energetic electrons gyrating in a strong magnetic field. But other mechanisms, such as thermal emission or energy loss by relativistic electrons in intense radiation fields, are also possibilities³. The radiation released through synchrotron emission is highly polarized, but the other suggested mechanisms do not naturally produce large polarization. That Coburn and Boggs detect a clear polarization in the γ -rays from a burst provides direct evidence in support of synchrotron emission as the mechanism of γ -ray production.

Their observations also reveal more about the nature of the magnetic field in which synchrotron emission occurs. In a GRB, the γ -rays are produced in different regions inside the jet (Fig. 2). These regions are unresolved by detectors close to the Earth, so the source appears to be point-like. The detected polarization signal is, then, an average over the polarization of radiation produced at different points within the source. If the direction of polarization varies randomly from place to place in the jet, then the observed polarization signal is likely to average out to zero. But this 'washing out' of the polarization signal will not happen if the polarization direction is the same everywhere. For polarization produced by the synchrotron mechanism, this means that the γ -ray-producing region is suffused by an ordered magnetic field, oriented in the same direction everywhere (Fig. 2a).

The direction of polarization reported by Coburn and Boggs² remained constant throughout the duration of the GRB, but the γ -ray flux varied significantly. So it seems unlikely that such a strong, constant, ordered field could be generated in the region where the γ -rays are produced. Rather, this suggests that the strong field originates near the collapsing object, and is then carried by (or perhaps even drives) the jet outwards from the source: the mechanism by which gravitational energy is extracted and powers the jet is, possibly, electromagnetic.

But there could be another explanation.

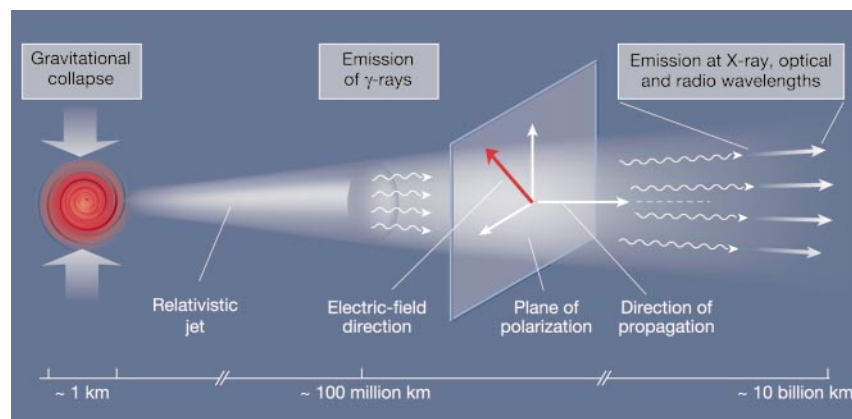


Figure 1 The creation of a γ -ray burst (GRB). If a dying star a few times heavier than our Sun collapses to a diameter of just one kilometre, a large amount of gravitational energy is released as a black hole forms. The energy is carried away by a highly relativistic jet of material, propagating at nearly the speed of light and generating, from inside the jet, a short flash of γ -rays. The electric field associated with the propagating γ -rays lies in a perpendicular plane and may point in any direction within this 'plane of polarization'. Coburn and Boggs² have detected a specific orientation, or polarization, of the electric field that could be the consequence of a strong, constant and well-ordered magnetic field in the region surrounding the source of the GRB. Further from the source, the jet interacts with the surrounding medium, generating an afterglow of X-rays, optical and radio waves that lasts a few days or months.

Strong polarization might also arise in a randomly oriented magnetic field — a structure that would be expected if the field is generated in the γ -ray production region — provided that the line-of-sight to the GRB lies close to the jet edge^{5,6} (Fig. 2b). In this case, the polarization signal is not averaged out: radiation reaches the observer only from points lying to one side of the line-of-sight, closer to the jet centre; radiation from points on the other side, outside the jet cone, is 'missing'. For a highly relativistic jet, with a velocity that is 99.99% that of light, such an orientation is likely to occur by chance only if its opening angle is close to 0.01 radians (0.6 degrees); the closer the jet velocity is to the speed of light, the smaller the opening angle required. Afterglow observations suggest that jet opening angles are typically about 0.05 radians, with more powerful GRBs produced by narrower

jets⁷. The burst reported by Coburn and Boggs² is exceptionally bright, and its jet may therefore have been very narrow. So the existence of a randomly oriented magnetic field cannot be discounted.

Both of these interpretations of the data pose challenges to models, which need to explain how a constant ordered field or a highly collimated jet might be produced. Future observations will determine which one is valid.

Eli Waxman is in the Faculty of Physics, Weizmann Institute of Science, Rehovot 76100, Israel.

e-mail: waxman@wicc.weizmann.ac.il

1. Fishman, G. J. & Meegan, C. A. *Annu. Rev. Astron. Astrophys.* **33**, 415–458 (1995).
2. Coburn, W. & Boggs, S. E. *Nature* **423**, 415–417 (2003).
3. Meszaros, P. *Annu. Rev. Astron. Astrophys.* **40**, 137–169 (2002).
4. Costa, E. et al. *Nature* **387**, 783–785 (1997).
5. Medvedev, M. V. & Loeb, A. *Astrophys. J.* **526**, 697–706 (1999).
6. Gruzinov, A. *Astrophys. J.* **525**, L29–L31 (1999).
7. Frail, D. A. et al. *Astrophys. J.* **562**, L55–L58 (2001).

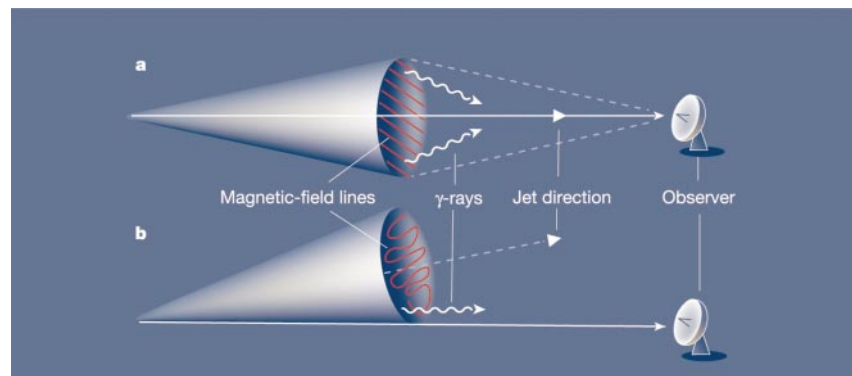


Figure 2 Magnetic fields and polarization. Energetic electrons gyrating in a strong magnetic field inside the jet of material ejected by a collapsing star would emit polarized γ -rays. The strong degree of polarization seen by Coburn and Boggs² suggests that the magnetic field is ordered, provided that the observer's line-of-sight to the γ -ray burst (GRB) is close to the axis of the jet cone (a). But if the line-of-sight to the GRB runs along the edge of the jet cone (b), the same degree of polarization could be seen even if the magnetic field is oriented randomly.



100 YEARS AGO

We learn from the *Athenaeum* that a Norwegian expedition, commanded by Captain Roald Amundsen, left Christiania a few days ago with the object of fixing the exact situation of the magnetic North Pole. The party are expected to be absent for four years, the route taken being by Lancaster Sound, Boothia Felix, where a magnetic observatory will be established for a period of two years under control of two members of the scientific staff, and back by the North-West Passage, Victoria Land, and the Behring Straits.

ALSO...

A Paris correspondent states that on May 8, a balloon built for MM. Lebaudy made a notable performance. The balloon left the Moisson Aërodrome in the morning and returned to it after having navigated round Mantes at a distance of 10 kilometres... The length of the air-ship is 56 metres, and the volume 2300 cubic metres. The engine is a 40 horse-power. There were two persons on board, M. Juchmès, a well-known professional aeronaut, and a mechanician.

From *Nature* 21 May 1903.

50 YEARS AGO

Sometimes, after heavy and prolonged onshore storms, great masses of foam are drifted in from the sea on to local rocks and beaches. A curious feature of these lather-like masses is the way in which they persist for relatively long periods even when blown about by the wind. Ordinary lathers soon revert to their former unlathered state: but sea foams may persist for a day or more when the weather continues moist and stormy and no sun shines. It is possible that the foam is 'held' by the presence of some protein material, and Miss E. M. Moore suggests that this protein material might be found in the alginate products of sea-weeds. This, however, does not explain how the surface tension of sea-water could be so lowered that the whipping action of storm waves could create a frothy mass... A likely explanation of the phenomenon is that, during storms, large numbers of planktonic organisms are destroyed and broken up because of the battering they receive in choppy seas. The surface tension-lowering chemicals which are thus released into the sea-water would allow the waves to whip up a froth and, if there is also present enough protein matter released from the alginate, a foam with lasting qualities might result.

From *Nature* 23 May 1953.

Gene expression

Silent clones speak up

Wolf Reik and Wendy Dean

An important category of genes — so-called pluripotency genes — are active in early embryos but silent in specialized cells. It seems that this silencing is difficult to reverse in cloned embryos.

Early on in the development of an embryo, the cells that are produced can contribute to a wide range of tissue types. But as development proceeds they become less and less 'pluripotent', as they are tailored towards specific functions. This restriction in developmental potential is associated with a relatively small number of genes being turned on at high levels, and a relatively large fraction of genes being switched off. A typical specialized cell thus expresses a minority of all the genes in the genome. An important question in both biology and medicine is whether and how these cells could be genetically 'reprogrammed' to adopt a different fate. The answers will provide insights into how cell specialization is maintained (which is of interest both fundamentally and in understanding cancer), and might open the way to personalized cell- and tissue-based treatments.

Reprogramming does occur during cloning: after the nucleus from a fetal or

adult cell is inserted into an egg, the genetic information in the nucleus can direct the development of a whole organism. But studies in *Genes and Development*¹ and *Development*² now show that, in cloned mouse embryos, the reactivation of genes that are silent in most adult tissues but are needed for early development is defective — perhaps explaining why cloning is inefficient³ (Fig. 1).

The two groups^{1,2} approached the issue of gene reprogramming in cloned embryos in a logical way. The genes they chose to study are repressed in adult somatic (non-reproductive) tissues but expressed in pluripotent ones (such as early embryos and primordial germ cells, which give rise to eggs or sperm). Moreover, at least one of these genes (*Oct4*) is actually necessary⁴ for the development of pluripotent cell lineages. Boiani *et al.*¹ looked specifically at *Oct4*, and found that a considerable proportion of cloned early embryos derived from cumulus cells (a somatic cell type frequently used for cloning) had

reduced levels or aberrant spatial patterns of *Oct4* expression. By contrast, many more early embryos that were produced by cloning with *Oct4*-expressing primordial germ cells continued to express this gene in the appropriate pattern. So, the reactivation of a somatically silenced gene seems to be difficult.

Bortvin *et al.*² went further, by also looking at ten genes that are related to *Oct4* in the sense of being expressed in similar patterns in pluripotent cells and repressed in differentiated ones. (The protein products of these genes are not necessarily related to each other, and their functions need to be determined.) The authors found that, with one exception (*Dppa5*), all these genes showed failures to be expressed in cloned embryos derived from cumulus cells. Notably, there was no relationship between which genes were re-expressed and which remained repressed — it is as if each gene has a certain probability of being reactivated, and the combination of these stochastic events determines overall expression patterns in individual clones. It is also remarkable that each of the 11 genes was re-expressed in at least one of the embryos studied, suggesting that the ability to reprogramme is present, at least in principle, but is perhaps limited. (Of course, the possibility cannot be excluded that this finding reflects a bias of the experimental set-up: perhaps embryos that did not activate a certain number of these genes died before the

RNA interference

Cereal adultery

For people who must restrict their protein intake — such as patients with kidney failure — a mutant rice that is naturally low in proteins called glutelins is beginning to be used as a dietary therapy. Makoto Kusaba and colleagues have now discovered how this mutant achieves low glutelin levels (*Plant Cell* doi:10.1105/tpc.011452; 2003). The answer involves the increasingly well-known biological phenomenon of RNA interference.

Glutelins are the major proteins in cereal grains such as wheat and rice. They are produced from two families of genes, the *GluA* and *GluB* families, which occur on at least three different chromosomes. Several mutations in rice disrupt one or another of these genes, but do not significantly reduce the total amount of glutelin produced. But the mutation studied by Kusaba *et al.* — 'low glutelin content-1', or LGC-1 — has a much broader effect, completely abolishing production of one *GluB* protein, radically reducing

the levels of other *GluB*s, and even limiting the *GluA* content.

Kusaba *et al.* have now found that LGC-1 plants lack a large stretch of DNA between two of the *GluB* genes, *GluB4* and *GluB5*. This region may encode a very short protein of about 25 amino acids, but loss of this protein does not appear to be the cause of the low glutelin levels. Crucially, the deletion also removes the 'stop' signal from the end of *GluB5*, so that the messenger RNA produced from this gene runs on into the *GluB4* mRNA. As it happens, these two genes have opposite orientations on the chromosome, and their sequences are almost completely complementary, so the mRNA folds over and the two sequences can bind to each other, forming double-stranded RNA.

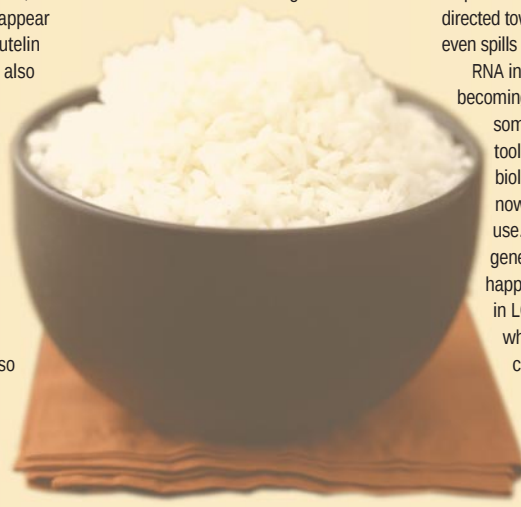
Double-stranded RNA is

involved in sequence-specific suppression of gene expression in organisms from plants to fungi to animals. The first step in this RNA-interference process is the production of smaller fragments, called small inhibitory RNAs, from a double-stranded RNA. These fragments act

as templates to guide suppression of specific genes by mechanisms including methylation. Kusaba *et al.* found potential small inhibitory RNAs, as well as increased methylation of glutelin genes, in LGC-1 plants. The similarity between glutelin gene sequences means that suppression is directed towards all *GluB* genes, and even spills over onto *GluA* genes.

RNA interference is fast becoming an invaluable, if sometimes unpredictable, tool in the molecular biologist's armoury. It has now found yet another use. The promiscuous gene silencer that happenstance has produced in LGC-1 plants might, when inserted as multiple copies into otherwise normal rice, result in still lower glutelin levels — a 'super-low-protein' rice.

Christopher Surridge



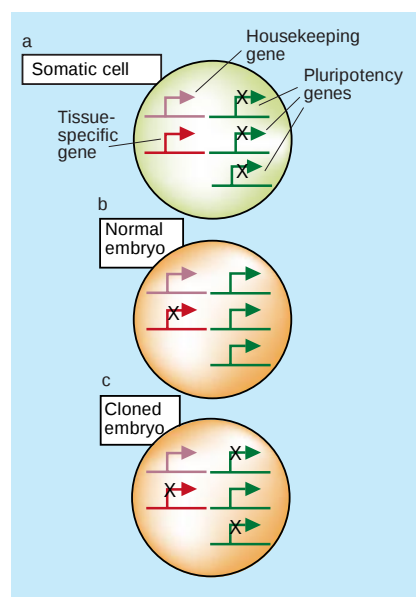


Figure 1 Why is cloning inefficient? The figure shows gene expression in specialized (somatic) tissues, normal embryos and cloned embryos. Housekeeping genes are needed for activities common to all cells; tissue-specific genes are activated only in particular somatic cells; and 'pluripotency' genes are expressed in embryonic cells that generate a wide range of tissue types. **a**, A somatic cell (such as a cumulus cell), with an active housekeeping gene and tissue-specific gene and three inactive pluripotency genes. **b**, In normal embryos, the housekeeping and pluripotency genes are expressed; the tissue-specific gene is repressed. **c**, As shown in the new studies^{1,2}, after cloning with the nucleus from a cumulus cell, gene activity is not reset correctly in many cloned embryos; in this example, whereas the housekeeping gene is expressed and the tissue-specific gene is silenced, only one of the pluripotency genes is reactivated.

stage at which the authors studied them.)

Bortvin *et al.*² also used embryonic stem (ES) cells, which express the 11 'pluripotency genes', for cloning, and found that expression continued in the cloned embryos. They argue that this might explain why ES-cell-derived clones are more likely to develop to term than cumulus-derived clones. What about genes that are normally expressed in somatic tissues and silent in early embryos? Bortvin *et al.* examined three such genes, and found that all were silenced in cloned embryos — perhaps implying that silencing is more efficient than reactivation. However, another study⁵ suggests that cloned embryos do retain some memory of the differentiated cells from which they were derived, in that a tissue-specific gene remained active, so this issue probably needs further investigation.

Several studies^{6–8} have now shown altered gene-expression profiles in cloned embryos, including amphibians. But these recent experiments^{1,2} that look at pluripotency

genes have taken our understanding further. They link gene-expression defects to defects in early development, and they provide good gene candidates with which to examine the precise mechanisms of reprogramming.

Why is reprogramming so difficult? The answer is probably that, once cells have differentiated into specific types, the silencing of unwanted gene expression is very tightly controlled, involving many reinforcing mechanisms. For instance, the modification of DNA with methyl groups (methylation) is commonly associated with gene silencing, as is the methylation of the histone proteins that bundle DNA into a compact form (chromatin) in the nucleus. These 'marking' mechanisms are likely to be connected in a way that makes the silent state very stable. Conversely, gene expression is often associated with histone acetylation.

It would be interesting to find out whether and how such marks can be reprogrammed, particularly on the genes studied by Boiani *et al.*¹ and Bortvin *et al.*². The early embryo can certainly reset the chromatin modifications characteristic of the male and female gametes from which it was formed^{9,10}. In terms of cloned embryos, so far only genome-wide chromatin reprogramming has been studied¹⁰. But it seems that, for the most part, the somatic patterns of histone methylation and acetylation are reset very inefficiently (although in a few embryos these marks look relatively normal, and are associated with a higher rate of successful development to a

crucial stage, the blastocyst stage). So chromatin modifications might indeed provide a mechanistic explanation for the difficulties in reactivating silent genes, and silencing active ones. Efficient reprogramming might require enzymes that remove acetyl groups from histones and methyl groups from DNA or histones (although DNA and histone 'demethylases' — if they exist at all — are still elusive¹¹).

Whatever the arguments for and against cloning, its study is already providing insight into the biology of cell differentiation, the extent to which cells can have many different fates, and the factors involved in reprogramming. With patience, this line of research should lead to more efficient and safer applications of reprogramming technologies in medicine. ■

Wolf Reik and Wendy Dean are in the Laboratory of Developmental Genetics and Imprinting, The Babraham Institute, Cambridge CB2 4AT, UK. e-mail: wolf.reik@bbsrc.ac.uk

1. Boiani, M., Eckardt, S., Schöler, H. R. & McLaughlin, K. J. *Genes Dev.* **16**, 1209–1219 (2002).
2. Bortvin, A. *et al.* *Development* **130**, 1673–1680 (2003).
3. Wilmut, I. *et al.* *Nature* **419**, 583–587 (2002).
4. Smith, A. G. *Annu. Rev. Cell. Dev. Biol.* **17**, 435–462 (2001).
5. Gao, S. *et al.* *Biol. Reprod.* doi:10.1095/biolreprod.102.014522 (2003).
6. Humpherys, D. *et al.* *Proc. Natl Acad. Sci. USA* **99**, 12889–12894 (2002).
7. Inoue, K. *et al.* *Science* **295**, 297 (2002).
8. Byrne, J. A., Simonsson, S. & Gurdon, J. B. *Proc. Natl Acad. Sci. USA* **99**, 6059–6063 (2002).
9. Santos, F., Hendrich, B., Reik, W. & Dean, W. *Dev. Biol.* **241**, 172–182 (2002).
10. Santos, F. *et al.* *Curr. Biol.* (in the press).
11. Bird, A. *Nature Immunol.* **4**, 208–209 (2003).

Condensed-matter physics

Thermopower to the people

Cronin B. Vining

The larger-than-expected thermally generated voltage seen in a layered-oxide material — which may prove useful in power generation or cooling — is now attributed to the spins of moving charges.

Thermocouples generate a voltage in a temperature gradient. This is known as 'thermopower', or the Seebeck effect, after its discoverer Thomas Johann Seebeck. These devices have found a range of applications, from cooling devices for seats in luxury automobiles to power supplies for spacecraft (including the Voyager missions; Fig. 1, overleaf). Metallic thermocouples generate relatively small voltages, but semiconductor thermocouples produce much larger voltages and can convert heat directly to electricity or generate cooling from an electrical input. Two different groups have reported semiconductor thermoelectric materials that are about twice as efficient as any previously known^{1,2}, achieved by carefully controlling the composition and structure of the materials on the atomic scale. But an entirely different approach to high thermopower

uses magnetic cobalt oxides — layered materials that combine the thermopower of semiconductors with the electrical conductivity of metals. On page 425 of this issue, Wang *et al.*³ account for their extraordinarily high thermopower.

These cobalt oxides ($\text{Na}_x\text{Co}_2\text{O}_4$) were first considered as thermoelectric materials by Terasaki *et al.*⁴ and have a variety of unusual properties. They are ionically bonded (unlike the classic semiconductor thermoelectric materials, which are covalently bonded), and can be doped with varying numbers of sodium atoms to achieve the desired properties. At room temperature, their thermopower is as much as ten times larger than might be expected, much larger than is typical of metals.

At the same time, $\text{Na}_x\text{Co}_2\text{O}_4$ has some unusual magnetic properties. At low

When these spins move, they carry some energy with them — and by doing so, contribute to the thermopower of the material. This doesn't happen in an ordinary metal, or in an ordinary magnet, because there the spins don't move.

The moving spins — or, more technically, spin entropy — are thought to be behind the enhanced thermopower of cobalt oxides. If this is the case, the thermopower should be suppressed if a magnetic field is applied in the plane of the layered-oxide material, blocking the motion of the spins. Wang *et al.*³ have shown, quite unambiguously, that this is indeed what happens (Fig. 4 on page 427): they offer a textbook-quality illustration of quantitative agreement between their experimental results and a remarkably simple (and classic) interpretation of the spin contribution to the thermopower, an agreement that holds over a wide range of temperatures

and magnetic fields. There are already reports that $\text{Na}_2\text{Co}_2\text{O}_4$ has thermoelectric properties comparable to the best known thermoelectric materials^{5,6} at temperatures near 1,000 K. Perhaps understanding the role that spin has to play in the thermopower of these oxides will enable further progress to be made.

Cronin B. Vining is at ZT Services, 2203 Johns Circle, Auburn, Alabama 36830-7113, USA.

e-mail: nature@zts.com

1. Venkatasubramanian, R., Silvola, E., Colpitts, T. & O'Quinn, B. *Nature* **413**, 597–602 (2001).
2. Harman, T. C., Taylor, P. J., Walsh, M. P. & LaForge, B. E. *Science* **297**, 2229–2232 (2002).
3. Wang, Y., Rogado, N. S., Cava, R. J. & Ong, N. P. *Nature* **423**, 425–428 (2003).
4. Terasaki, I., Sasago, Y. & Uchinokura, K. *Phys. Rev. B* **56**, 12685–12687 (1997).
5. Fujita, K., Mochida, T. & Nakamura, K. *Jpn. J. Appl. Phys.* **40**, 4644–4647 (2001).
6. Ohtaki, M., Nojiri, Y. & Maeda, E. in *Proc. 19th Int. Conf. Thermoelectrics* (ed. Rowe, D. M.) 190–195 (Babrow, Cardiff, 2000).

Alzheimer's disease

Mental plaque removal

Bart De Strooper and James Woodgett

Lithium, already used to treat psychiatric disorders, has been found to reduce amyloid-peptide production in a mouse model of Alzheimer's disease. The implication is that lithium's target molecule helps to generate the peptides.

Studies of people with Alzheimer's disease have revealed several notable changes in the brain, including the build-up of protein deposits known as amyloid plaques outside nerve cells, and the accumulation of so-called neurofibrillary tangles inside the cells. Should treatments for this disease target the plaques or the tangles? On page 435 of this issue, Phiel and co-workers¹ propose the idea of tackling both with a single drug — one that inhibits the enzyme glycogen synthase kinase-3 (GSK-3).

This enzyme is an evolutionarily conserved serine/threonine kinase that, in response to cellular signalling events, modifies the functions of a variety of proteins by specifically adding phosphate groups to the amino acids serine or threonine². It is one of the candidates for producing excessive phosphorylation of the tau protein — and 'hyperphosphorylated' tau is the main component of the neuronal tangles seen in patients with Alzheimer's disease. The precise effects of these tangles are unknown, and indeed it remains a matter of debate whether the phosphorylation of tau by GSK-3 *in vivo* harms³ or protects⁴ neurons. Nonetheless, the accumulation of tangles does correlate with progression of the disease.

So, too, does the build-up of amyloid plaques. One of the major components of these plaques is a small peptide known as amyloid- β (A β) peptide, which is produced from a longer protein — the amyloid precursor

protein (APP) — that sits in neuronal membranes. APP is precisely snipped in two places to generate A β . The first cleavage is carried out by an enzyme called β -secretase, or by metalloprotease enzymes⁵. This cleavage generates a membrane-bound fragment of APP (the carboxy-terminal fragment), which is then severed by the γ -secretase complex. The A β peptide is thereby produced, and is shed from the neuronal surface (Fig. 1). The prevalent hypothesis is that A β peptides, congregating outside nerve cells in the brain, are the main trigger for the neuronal degeneration characteristic of Alzheimer's disease⁶.

The past decade has seen strong investment in the development of drugs that tackle the deposition of amyloid plaques, using various approaches⁷. On the other side of the coin, the idea of using GSK-3 inhibitors to prevent the build-up of tangles has been around for some time (although validation of this strategy has been problematic, mainly because there are still no good animal models of tangle accumulation, but also because of the question marks over the importance of the *in vivo* phosphorylation of tau by GSK-3). Phiel *et al.*¹ now add an interesting twist to the issue. They find that lithium and kenpaullone — two structurally unrelated inhibitors of GSK-3 — interfere with the production of A β peptides in cell culture. This has been found previously for lithium and with different cells⁸, but Phiel *et al.* go further

Figure 1 Powerful heat. In 1822, the Estonian–German physicist Thomas Johann Seebeck (inset) discovered that if heat is applied across the junction of two wires, a current is generated. Seebeck called the effect 'thermomagnetism' (a term later replaced by the more accurate 'thermoelectricity'). It is the basis of thermocouples — devices used for cooling, and for power generation such as in the Voyager mission to Jupiter and Saturn, seen here at its launch in 1977.

temperatures it is an antiferromagnet, which means that there are 'spins' in the material that take on a particular kind of order. Unlike most magnets, the spins in $\text{Na}_2\text{Co}_2\text{O}_4$ are not fixed to specific atoms within the lattice but instead are free to move about the crystal.

and show that the inhibitors have the same effect *in vivo* in a mouse model of Alzheimer's disease. So GSK-3 inhibitors could be an attractive addition to the list of candidate drugs for preventing amyloid plaques — with the potential benefit of inhibiting the formation of tangles, too (Fig. 1).

Several questions need answering first, however. One issue is the mechanism by which lithium and kenpaullone inhibit plaque formation. Phiel *et al.* show that membrane-bound carboxy-terminal fragments of APP accumulate in the presence of GSK-3 inhibitors. A similar build-up is observed when the γ -secretase complex is genetically or pharmacologically inactivated⁹ — so a GSK-3-dependent process might normally be required for γ -secretase activity. But how? In principle, any of the four proteins known to make up the γ -secretase complex (presenilin-1, APH-1, PEN-2 and nicastrin¹⁰) could be a target for GSK-3. For instance, presenilin-1 binds both to GSK-3 and to one of its best known substrates, β -catenin^{11,12}. And presenilin-1 itself is phosphorylated by this enzyme¹³. The more likely explanation, however, is that GSK-3 regulates the cleavage of APP at the level of the carboxy-terminal fragments, not the γ -secretase complex. In support of this, Phiel *et al.* find that lithium does not affect the cleavage of Notch — another well-established substrate of γ -secretase that is involved in cell-fate decisions in embryos and adults.

Another question concerns the precise effects of the two highly related isoforms of GSK-3 — GSK-3 α and GSK-3 β — in the build-up of amyloid plaques. Relatively little is known about the respective roles of these two kinases in other biological situations. The proteins are encoded by distinct genes, and although their catalytic sites are well conserved, they differ in their amino- and carboxy-terminal parts². Mouse embryos lacking GSK-3 β develop relatively normally until day 12, indicating that GSK-3 α can step into GSK-3 β 's role in the developmentally important β -catenin pathway¹⁴. But it cannot totally compensate for its relative: for example, later in development, it can't fill GSK-3 β 's role in protecting the liver from degeneration¹⁴.

To assess whether GSK-3 α , GSK-3 β or both are important in the production of amyloid plaques, Phiel *et al.*¹ used RNA-mediated gene silencing to decrease the expression of each isoform in turn. Although the silencing was incomplete, the data are consistent with a positive role for GSK-3 α , and a possible negative role for GSK-3 β , in the production of A β peptides. Moreover, lithium inhibits both isoforms, and its overall effect is to block A β generation, suggesting that the role of GSK-3 β in this regard is subservient to that of GSK-3 α . That, however, remains to be shown conclusively.

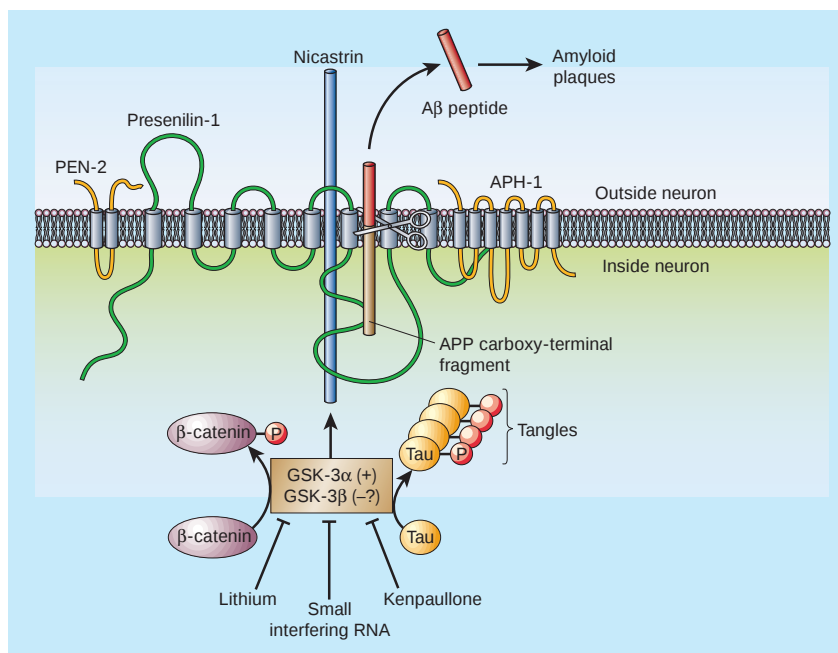


Figure 1 Plaques and tangles in Alzheimer's disease. Amyloid plaques contain the A β peptide, which is produced from the amyloid precursor protein (APP). Cleavage of APP at one point (not shown) generates a carboxy-terminal fragment. This is then severed by the γ -secretase complex (which consists of at least four proteins¹⁰ — presenilin-1, APH-1, PEN-2 and nicastrin), producing A β . Tangles are produced from hyperphosphorylated tau protein, possibly by means of the enzyme glycogen synthase kinase-3 (GSK-3). GSK-3 also regulates phosphorylation of β -catenin (which can bind presenilin-1). Phiel *et al.*¹ find that GSK-3 might also regulate — either directly or indirectly — the cleavage of APP carboxy-terminal fragments. They show that lithium and kenpaullone (two GSK-3 inhibitors), as well as small interfering RNAs that prevent GSK-3 expression, inhibit A β production. So, GSK-3 inhibitors might interfere with the generation of both amyloid plaques and tangles in Alzheimer's disease. The precise roles of the two forms of GSK-3, α and β , are unknown.

Part of the excitement of Phiel and colleagues' work comes from the fact that lithium affects the cleavage of APP without appearing to interfere with Notch processing. Until now, nearly all γ -secretase inhibitors have affected both processes equally¹⁵, which has been a problem given Notch's importance. On the downside, however, the inhibition of GSK-3 is fraught with the possibility of severe side effects, such as the induction of intestinal cancer through effects on β -catenin. Moreover, lithium — a long-established treatment for psychiatric disorders, particularly bipolar disorder — is poorly tolerated by elderly patients, and harms the kidneys at high concentrations, making the therapeutic window very small. Luckily, however, the effective concentrations of lithium in cell culture, and the serum lithium concentrations in mice, lie in the clinically acceptable range¹, suggesting that many potential patients may benefit.

An obvious question now is whether the use of lithium or other GSK-3 inhibitors correlates negatively with dementia. This will probably be difficult to determine from long-term follow-up studies of patients who are taking this drug as a treatment for bipolar disorders. So a more focused clinical trial will

be needed to decide whether lithium — and, later, more selective GSK-3 inhibitors — becomes a new weapon in the arsenal of drugs to treat Alzheimer's disease. ■

Bart De Strooper is at the Center for Human Genetics, Flanders Interuniversity Institute for Biotechnology, and the KU Leuven, Herestraat 49, 3000 Leuven, Belgium.

e-mail: bart.destrooper@med.kuleuven.ac.be

James Woodgett is at the Ontario Cancer Institute, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada.

e-mail: jwoodgett@uhnres.utoronto.ca

- Phiel, C. J., Wilson, C. A., Lee, V. M.-Y. & Klein, P. S. *Nature* **423**, 435–439 (2003).
- Doble, B. W. & Woodgett, J. R. *J. Cell Sci.* **116**, 1175–1186 (2003).
- Lucas, J. J. *et al.* *EMBO J.* **20**, 27–39 (2001).
- Spittaels, K. *et al.* *J. Biol. Chem.* **275**, 41340–41349 (2000).
- Annaert, W. & De Strooper, B. *Annu. Rev. Cell. Dev. Biol.* **18**, 25–51 (2002).
- Hardy, J. & Selkoe, D. J. *Science* **297**, 353–356 (2002).
- Dominguez, D. I. & De Strooper, B. *Trends Pharmacol. Sci.* **23**, 324–330 (2002).
- Sun, X. *et al.* *Neurosci. Lett.* **321**, 61–64 (2002).
- De Strooper, B. *et al.* *Nature* **391**, 387–390 (1998).
- De Strooper, B. *Neuron* **38**, 9–12 (2003).
- Takashima, A. *et al.* *Proc. Natl. Acad. Sci. USA* **95**, 9637–9641 (1998).
- Kang, D. *et al.* *Cell* **110**, 751–762 (2002).
- Kirschbaum, F., Hsu, S. C., Cordell, B. & McCarthy, J. V. *J. Biol. Chem.* **276**, 7366–7375 (2001).
- Hoeftlich, K. P. *et al.* *Nature* **406**, 86–90 (2000).
- De Strooper, B. *et al.* *Nature* **398**, 518–522 (1999).

Synthetic chemistry

A perfect fit

V. Ramamurthy

Two molecules may form a dimer if lured into an organic 'nanocage'. So far, the approach has only worked for identical molecules. But two different molecules will dimerize if, together, they snugly fit the dimensions of the cage.

The inspiration of nature has led chemists to investigate confinement as a means of achieving selectivity in reactions: from the remarkable selectivities exhibited by enzymes in diverse thermal and photochemical processes, to the exquisite phenomenon of photosynthesis, a confined environment has the power to drive a reaction. Intrinsic reactivity frequently becomes of secondary importance compared with features such as site symmetry, nearest-neighbour separation and other geometric considerations. In a confined environment, the familiar electronic and steric effects of solution chemistry are replaced by structural and topological factors that frequently result in products that display a level of selectivity or stereospecificity that has hitherto been unobtainable in solution¹.

Notable among the many ordered or constrained media that have been investigated as selective environments are micelles, micro-emulsions, liquid crystals and solid phases — particularly porous solids such as silica, alumina, clay and zeolites². But, despite considerable progress using such media, the important problem of localizing and aligning two different types of molecules has remained. Now Makoto Fujita's group, writing in the *Journal of the American Chemical Society* as Yoshizawa *et al.*³, outline their success in bringing two different molecules within reactive distance by using a self-assembled confined environment — a nanocage.

In earlier work, Fujita and colleagues demonstrated the assembly, in water, of a nanocage consisting of six palladium ions and four tridentate ligands⁴ (Fig. 1). The empty cage, with an interior diameter of 8 Å, can accommodate various hydrophobic guest molecules. Localizing reactive molecules in hydrophobic cages is not unprecedented, but Fujita and co-workers realized that trapping two different molecules (say, A and B) in a single cage would work best if the other possible combinations (A and A, and B and B) were either too large or too loose a fit. Selective inclusion would then be driven by the size compatibility of the host and its guests.

Fujita and co-workers have explored³ this idea by investigating the light-initiated cycloaddition of acenaphthylene (molecule A) to 5-ethoxynaphthoquinone (molecule B) and to *N*-benzylmaleimide (molecule C). In the absence of the nanocage, irradiation of a mixture of A and B or A and C in benzene gave only the acenaphthylene homodimer (AA); neither the heterodimer (AB or AC) nor the homodimer of B or C was obtained. On the other hand, when mixtures of A and B or A and C in water were irradiated in the presence of the nanocage, heterodimers did form. A small change in the structure of one of the reactants (for example, replacing an ethyl group on molecule B with a methyl group, or replacing the benzyl group on molecule C with a methyl group) led to loss of selectivity for the heterodimeric products. Such slight

structural changes would not have influenced the product distribution if, say, a charge-transfer interaction had been responsible for bringing the reactants together. So these results are consistent with the authors' claim that the selectivity is due to size compatibility.

Previous attempts to organize two different reactant molecules in a confined environment have met with limited success. The earliest work attempted mixed photodimerization of *trans*-cinnamic acids and *trans*-cinnamides by irradiating co-crystals of the two species⁵ — but the results of this approach are unpredictable. The alignment of two different molecules (3-alkylcyclopentenone and 1-heptenyl acetate) at the interface of micelles followed by photodimerization had some success⁶, but although this approach is more predictable, product selectivity is poor. And in an investigation of the Diels–Alder ring-forming addition reaction in natural (cyclodextrin) and synthetic cages^{7,8}, it remained unclear whether all the cages actually contained two different molecules.

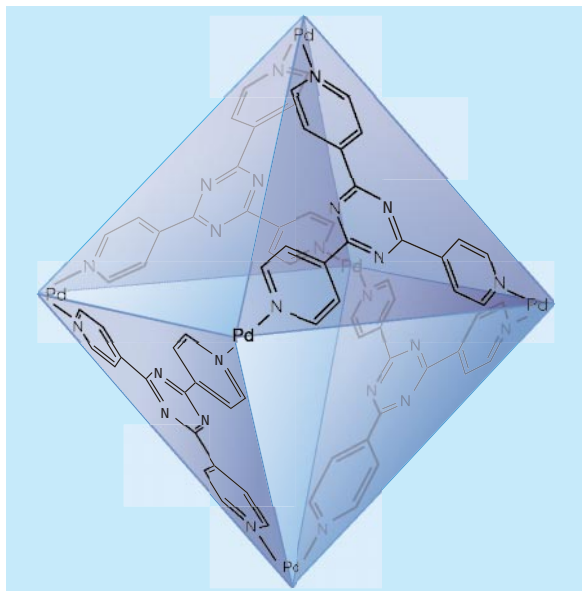
Admittedly, the method described by Fujita and co-workers has limited generality, but it ensures that all cages contain two different molecules within reactive distance. The commercial availability of nanocages and the easy inclusion of guest molecules (by suspending them and the cages in water at 80 °C for about ten minutes) make this approach appealing. An inventory should be made of the dimerization processes possible in cages of several sizes.

Most chemical phenomena — energy transfer, electron transfer, molecular recognition and bimolecular reactions, to name just a few — require specific communication between two different molecules, and for that, the two interacting molecules must be arranged in a fixed geometry close to each other. Often this is achieved with the help of the strong intermolecular forces that are associated with charge transfer, multiple hydrogen bonding or dipolar and quadrupolar interactions. But such factors limit the approach to molecules that contain the specific functionalities to respond to these forces. The nanocage method based on the complementarity of molecular morphologies, however, has limitless potential. ■

V. Ramamurthy is in the Department of Chemistry, Tulane University, New Orleans, Louisiana 70118, USA.

e-mail: murthy@tulane.edu

Figure 1 The nanocage. This self-assembled structure of organic molecules, with palladium (Pd) ions at the vertices, has an interior diameter of about 8 Å (large enough to hold four benzene molecules). In aqueous solution, hydrophobic molecules are lured into the cage. Fujita and colleagues have shown³ that if the combined volume of two different molecules is a close match to the dimensions of the cage, the two species dimerize easily upon irradiation.



- Weiss, R. G., Ramamurthy, V. & Hammond, G. S. *Acc. Chem. Res.* **26**, 530–536 (1993).
- Ramamurthy, V. (ed.) *Photochemistry in Organized and Confined Media* (VCH, New York, 1991).
- Yoshizawa, M., Takeyama, Y., Okano, T. & Fujita, M. *J. Am. Chem. Soc.* **125**, 3243–3247 (2003).
- Fujita, M. *et al. Nature* **378**, 469–471 (1995).
- Hung, J. D., Lahav, M., Luwisch, M. & Schmidt, G. M. I. *Isr. J. Chem.* **10**, 585–599 (1972).
- de Mayo, P. & Sydnes, L. K. *J. Chem. Soc. Chem. Commun.* 994–995 (1980).
- Kang, J. & Rebek, J. *Nature* **385**, 50–52 (1997).
- Breslow, R. & Guo, T. *J. Am. Chem. Soc.* **110**, 5613–5617 (1988).

Vision

Sea gypsies see shells on the sea floor

Curr. Biol. **13**, 833–836 (2003)

Under water, most of us are half-blind — but not so the Moken. This Southeast Asian tribe of sea gypsies can see under water twice as clearly as Europeans, according to Anna Gislén and colleagues.

The semi-nomadic Moken, who live among the Surin Islands off Thailand's west coast, use their superior visual skills to dive for food on the ocean floor. Most of us see blurred images when we dive without goggles or a mask, as our eyes struggle to bend and focus light under water. But Moken children can pick out small shells, clams and sea cucumbers at depths of 3 to 4 metres.

Gislén *et al.* have now compared the sub-aqua vision of native Moken and holidaying European children. Immersed in the clear, shallow waters of the Ko Surin National Park, the Moken children were able to distinguish underwater objects less than 1.5 millimetres wide. Europeans struggled to make out anything less than 3 millimetres across.

The authors found that, unlike the visitors, the Moken swimmers constrict their pupils while diving. The researchers also speculate that the Moken can squeeze the lenses of their eyes more, making them thicker and better able to bend incoming light. Together, the two processes bring blurry images into sharper focus, pushing the eyes' optics to the limit of what may be humanly possible.

Helen R. Pilcher

Biogeochemistry

Heavy metal out of air

Atmos. Environ. **37**, 1613–1622 (2003)

Do deciduous trees simply recycle mercury in the environment by taking it up from the soil and then returning it when the leaves drop off in autumn? Or do leaves capture gaseous mercury from the atmosphere, and when they fall add to the local soil concentrations?

J. A. Erickson and colleagues investigated these questions in a two-year study of quaking aspen trees (*Populus tremuloides*), grown under highly controlled conditions. They conclude that almost all of the mercury in leaf tissue comes from the atmosphere, and that its passage through foliage into the terrestrial environment probably constitutes a significant sink for the element.

Mercury is a serious pollutant, and there is a lot of it about. Earlier this year, the United Nations Environment Programme estimated that emissions have tripled since pre-industrial times, and that some

2,000 tonnes are released annually from coal-burning power stations and mining operations (in which mercury is used in precious-metal extraction).

The methylated form of the element is especially nasty. In provisional results, Erickson *et al.* find that the amount of methylmercury in foliage decreases with leaf age. But how mercury is transformed during its passage from atmosphere to plant to soil and beyond remains largely unknown.

Tim Lincoln

Microbiology

Flab control

Dev. Cell **4**, 663–672 (2003)

Fat does have its uses. Lipids are an essential component of cell membranes and a source of regulators of metabolism and energy. Their production is complex and tightly regulated, so each membrane can be tailored to suit the needs of the cell, the tissue and the organism. Gustavo E. Schujman and colleagues have now found a molecular switch that controls lipid biosynthesis in the bacterium *Bacillus subtilis*.

Schujman *et al.* discovered that one *B. subtilis* protein, encoded by the *fapR* gene, could bind to and repress several clusters of genes involved in the production of fatty acids. Deleting the *fapR* gene turned on all these 'fat-production genes' at once. This slowed down the growth of the bacterium and also made it less resistant to cold — probably because of alterations in the lipid composition of its membrane.

In *B. subtilis* that did carry *fapR*, however, the fat-production genes could be switched on by inhibiting fatty-acid biosynthesis. So, the FapR protein seems able to gauge the amount of fatty-acid synthesis required. When fatty-acid levels fall below the norm, FapR relieves the fat-production genes from repression. And, the authors propose, it may be the build-up of biosynthetic intermediates that removes FapR from these genes, and kick-starts the fatty-acid production cycle.

Marie-Thérèse Heemels

Behavioural ecology

Fresh with the tang of citrus

Proc. R. Soc. Lond. B doi:10.1098/rspb.2003.2379 (2003)

Seabird colonies usually smell rank. But the crested auklet (*Aethia cristatella*, pictured) adds a more pleasant note to the guano — the scent of tangerines. Julie C. Hagelin and co-workers say their research shows that these are the first birds found to use smell as a social signal.

It's not yet clear what that signal means. Hagelin *et al.* speculate that it could be an olfactory ornament: perhaps only well-fed birds can produce lots of citrusy perfume, showing themselves to be good mates.



Crested auklets — signalling with scent.

Males and females are equally pungent, and monogamous. Both sexes are picky about who they mate with.

The birds live in large colonies on the coast of Alaska. Given a choice between their citrus smell, on either feathers or chemical-soaked cotton wool, and an animal odour or a bland control, Hagelin and colleagues found that auklets spent much more time investigating their signature scent. When two auklets meet, they bury their bills in the feathers around each other's necks, where the smelliest feathers are found.

Birds have a perfectly good sense of smell, which they use in finding food and navigating. The apparent absence of odour signals has been a puzzle, but researchers may now be prompted to sniff out more avian chemical messages.

John Whitfield

Fluid dynamics

How to crack a liquid

Phys. Rev. Lett. **90**, 184501 (2003)

Dribble water from a hose into a swimming pool, and the liquid mixes smoothly. Turn up the flow to make a jet, and the impact creates bubbles. Why the change, ask Élise Lorenceau and colleagues? That question can be important in industrial processes such as pouring liquids into moulds, where bubbling creates flaws in the cast objects.

Lorenceau *et al.* study the problem using a cylinder rotating horizontally at the interface between a dense fluid (silicone oil, glycerol) and a lighter fluid (usually, in fact, air). A film of the dense fluid coats the cylinder surface, is carried around through the air, and is pushed down, like a jet, back into the body of dense fluid. This creates a cusp-like indentation of air along the downward-moving edge.

When does this cusp break up to cause entrainment of air? Theory and experiment show that the cusp's tip gets sharper as the rotation velocity increases. At some critical speed, the cusp becomes unstable: the denser fluid 'fractures', and air is pulled down in a narrow film, which eventually breaks up into bubbles. This critical velocity depends mainly on the viscosity of the denser fluid, but also on that of the lighter one.

Philip Ball

Rapid change in mouse mitochondrial DNA

Wild mice around Chicago may have switched genotype to keep pace with modern living.

We have compared the sequences of mitochondrial DNA extracted from museum skins of white-footed mice caught in the Chicago area since 1855 and from modern mice trapped alive in the same locations. We found a consistently similar directional change of mouse genotype over this period at each of five collection sites that were separated by 10–70 km. The genotype most common 100 years ago is now extremely rare, indicating that the mammalian mitochondrial genome can undergo rapid evolution.

Museums from Alaska to Switzerland were searched for specimens of the white-footed mouse, *Peromyscus leucopus*, originally collected from two counties in northeastern Illinois, USA. The white-footed mouse is the most common rodent in the deciduous forests of North America. We borrowed 61 museum-specimen skins of mice collected from five locations in the Chicago area, and in 1999 we trapped 52 white-footed mice in the same locations and in three others.

Small (12 mm × 1.5 mm) strips were cut from the ventral suture of the museum-specimen mouse skins. We found that shaving the hair from these strips before extraction, combined with a freeze-thaw step, reduced the dark colour of the extract and prevented inhibition of the polymerase chain reaction (PCR)¹. The Chelex-100 DNA-extraction protocol² was modified by the addition of proteinase K (ref. 3).

We used a 340-base-pair polymorphic region within the coding portion of the cytochrome oxidase II gene to genotype the museum mouse skins. This region was amplified for sequencing of DNA from 56 of 61 skins by using KlenTaqLA DNA polymerase and PCR buffer (DNA Polymerase Technology)^{4–7}, extremely long (20-min) extension cycles⁵ and betaine⁸, in conjunction with three nested PCR primer sets spanning 559, 496 and 400 base pairs in turn.

Three haplotypes (named A, M and Mw) were identified: Mw was found in only three mice. M and Mw differ at just a single position, but they both differ from A at four positions in this 340-base-pair region (see supplementary information); none of the base changes alters the amino-acid sequence of the gene.

At four of the five collection sites, the oldest white-footed mice were predominantly of genotype A. At each geographical location there was a monotonic decrease in the proportion of A (Fig. 1). Pooling sequencing data for all five locations, the proportion of A was 5/5 in 1850–99, 18/30

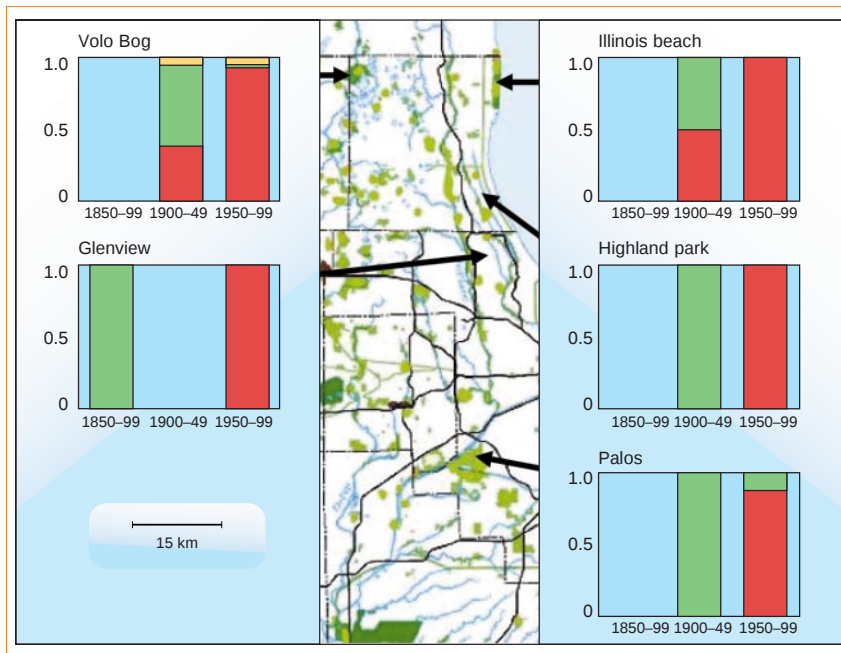


Figure 1 Modern map of the Chicago region where the change in genetic diversity of white-footed mice over 150 years was analysed. The geographic relation is shown between the five locations (arrows) from which pre-1950 museum specimens of *Peromyscus leucopus* were obtained. Major highways are shown in black; natural areas are shown in green. The frequencies (vertical scale) of three different mitochondrial haplotypes (A, green; M, red; Mw, yellow) are shown for each location for the periods 1850–99, 1900–49 and 1950–2000.

in 1900–49 and 4/73 in 1950–99. The observed frequency change requires less than a 1% per generation advantage of M over A. None of the live mice trapped at these five locations had the A type, and only one A mouse was trapped in 1999–2000.

As the change in the mitochondrial genotype of the wild mouse coincides with a marked increase in human activity in the region, we presume that this caused the replacement of the A haplotype by the M. The M genotype might have become advantageous in the altered habitat, or it could be unconditionally advantageous and have been introduced by people from elsewhere.

In the Chicago region, the white-footed mouse (Fig. 2) has displaced the prairie deer mouse from the few remaining

prairies^{9,10}. We suggest that the M haplotype has not only spread through the white-footed mouse population but might also have contributed to the displacement of the prairie deer mouse by the white-footed mouse.

Oliver R.W. Pergams*†, Wayne M. Barnes‡, Dennis Nyberg*

*Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois 60607, USA

†Present address: Department of Conservation Biology, Chicago Zoological Society and Smith Fellows Program, The Nature Conservancy, Brookfield, Illinois 60513, USA
e-mail: olpergam@brookfieldzoo.org

‡Department of Biochemistry and Molecular Biophysics, Washington University in St Louis School of Medicine, St Louis, Missouri 63110, USA



Figure 2 The white-footed mouse, originally a forest-dweller.

- Goodyear, P. D., MacLaughlin-Black, S. & Mason, I. J. *Biotechniques* **16**, 232–233 (1994).
 - Walsh, P. S., Metzger, D. A. & Higuchi, R. *Biotechniques* **10**, 506–513 (1991).
 - Steinberg, E. K. *Mol. Ecol.* **8**, 1075–1092 (1999).
 - Barnes, W. M. *Gene* **112**, 29–35 (1992).
 - Barnes, W. M. *Trends Biochem. Sci.* **19**, 342 (1994).
 - Barnes, W. M. *Proc. Natl Acad. Sci. USA* **91**, 2216–2220 (1994).
 - Barnes, W. M. US Patent No. 5,436,149 (1995).
 - Baskaran, N. et al. *Genome Res.* **6**, 633–638 (1996).
 - Pergams, O. R. W. & Nyberg, D. J. *Mammalogy* **82**, 984–992 (2001).
 - Pergams, O. R. W. & Nyberg, D. *Proc. R. Soc. Lond. B* (submitted).
- Supplementary information** accompanies this communication on Nature's website.
Competing financial interests: declared (see online version).

High-temperature superconductors

Universal nodal Fermi velocity

The mechanism that causes high-temperature superconductivity in copper oxide materials (cuprates) is still unknown, more than 15 years after it was discovered¹. As the charge carriers (electrons or holes) are introduced into the parent antiferromagnetic insulator, a process called doping, the material evolves from an insulator to a superconductor, and eventually to a normal metal. This marked change of physical properties with doping^{2–7} indicates that doping dependence (non-universality) might be a general feature of these materials, but we find that, on the contrary, the low-energy Fermi velocity of electrons is in fact universal, even among different superconductor families.

We used high-resolution angle-resolved photoemission, which can directly probe how electrons move in materials⁸, to investigate hole-doped materials at various dopings in five different families of cuprates. These included $(\text{La}_{2-x}\text{Sr}_x)\text{CuO}_4$ (LSCO) and $(\text{La}_{2-x-y}\text{Nd}_y\text{Sr}_x)\text{CuO}_4$ (Nd-LSCO), $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ (Bi2212), $\text{Bi}_2\text{Sr}_2\text{CuO}_6$ (Bi2201), $(\text{Ca}_{2-x}\text{Na}_x)\text{CuO}_2\text{Cl}_2$ (Na-CCOC) and $\text{Tl}_2\text{Ba}_2\text{CuO}_6$ (Tl2201). The LSCO system in particular covers the entire doping range ($0 < x \leq 0.3$) over which the physical properties vary from insulator ($0 \leq x < 0.03$) to superconductors ($0.05 < x < 0.25$) to overdoped non-superconducting metal ($x > 0.25$). Apart from the Na-CCOC data taken at the Stanford Synchrotron Radiation Laboratory, all samples were measured

at the Advanced Light Source at Lawrence Berkeley National Laboratory (experimental details are presented elsewhere^{9,10}).

Figure 1a shows the energy–momentum (dispersion) curves of the LSCO system with various dopings ($0 < x \leq 0.3$), measured along the $(0,0) - (\pi,\pi)$ direction in the Brillouin zone in reciprocal space. This diagonal direction is special in these superconductors because the anisotropic superconducting gap¹¹, as well as the normal-state pseudogap^{12,13}, is zero along this so-called nodal direction. Figure 1a shows that, for all dopings, there is a kink at an energy of about 70 meV that separates the dispersion into a high-energy part (that is, further from the Fermi energy) and a low-energy part (that is, closer to the Fermi energy) with different slopes. Whereas the high-energy dispersion varies with doping, the dispersion converges within about 50 meV of the Fermi energy, revealing a behaviour that is independent of doping. Correspondingly, a decrease is seen in the electron-scattering rate at an energy of about 70 meV, as indicated in Fig. 1b for the LSCO ($x = 0.063$) sample.

The electron velocity can be extracted quantitatively from the slope in dispersion, as $v = \partial\epsilon/\partial k$. We have obtained the low-energy velocity (the Fermi velocity) and the high-energy velocity as a function of doping for all five families of materials (see supplementary information). The Fermi velocity is nearly constant for all materials and dopings within an experimental error of about 20%. In contrast, the high-energy velocity varies strongly with doping.

This invariance of nodal Fermi velocity in cuprates is surprising, given the range

of variation in many other physical properties^{2–7}. This universal behaviour, together with the ubiquitous existence of a kink in the dispersion and a decrease in the scattering rate, are puzzles that require answers before the mystery of high-temperature superconductivity can be solved.

X. J. Zhou^{1,2} et al.*

e-mail: xjzhou@lbl.gov

*T. Yoshida^{1,3}, A. Lanzara^{1,2}, P. V. Bogdanov¹, S. A. Kellar¹,

K. M. Shen¹, W. L. Yang^{1,2}, F. Ronning¹, T. Sasagawa¹,

T. Kakeshita⁴, T. Noda⁴, H. Eisaki⁴, S. Uchida⁴, C. T. Lin⁵,

F. Zhou⁶, J. W. Xiong⁶, W. X. Ti⁶, Z. X. Zhao⁶, A. Fujimori^{1,2},

Z. Hussain⁷, Z.-X. Shen¹

¹Department of Physics, Applied Physics and SSRL, Stanford University, Stanford, California 94305, USA; ²Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA; ³Departments of ⁴Physics and

⁵Superconductivity, University of Tokyo, Bunkyo-ku, Tokyo 113,

Japan; ⁶Max-Planck-Institut für Festkörperforschung, 70569 Stuttgart,

Germany; ⁷National Laboratory for Superconductivity, Institute of

Physics, Chinese Academy of Sciences, Beijing 100080, China.

1. Bednorz, J. G. & Müller, K. A. *Z. Phys. B* **64**, 189–193 (1986).

2. Ando, Y., Lavrov, A. N., Komiya, S., Segawa, K. & Sun, X. F.

Phys. Rev. Lett. **87**, 17001–17004 (2001).

3. Uchida, S. et al. *Phys. Rev. B* **43**, 7942–7954 (1991).

4. Yamada, K. et al. *Phys. Rev. B* **57**, 6165–6172 (1998).

5. Loram, J. W., Luo, J. L., Cooper, J. R., Liang, W. Y. & Tallon, J. L. *Physica C* **341–348**, 831–834 (2000).

6. Uemura, Y. J. et al. *Phys. Rev. Lett.* **62**, 2317–2320 (1989).

7. Timusk, T. & Statt, B. *Rep. Prog. Phys.* **62**, 61–122 (1999).

8. Huefner, S. *Photoemission Spectroscopy: Principles and Applications* (Springer, Berlin, 1995).

9. Zhou, X. J. et al. *Phys. Rev. Lett.* **86**, 5578–5581 (2001).

10. Lanzara, A. et al. *Nature* **412**, 510–514 (2001).

11. Shen, Z.-X. et al. *Phys. Rev. Lett.* **70**, 1553–1556 (1993).

12. Loeser, A. G. et al. *Science* **273**, 325–329 (1996).

13. Ding, H. et al. *Nature* **382**, 51–54 (1996).

Supplementary information accompanies this communication on

Nature's website.

Competing interests: declared none.

Marine ecology

Spring algal bloom and larval fish survival

The different factors that influence the prevalent decline in fish stocks are currently subject to urgent and intense scrutiny. Here we combine the use of remote-sensing satellite data with a long-term data set of haddock recruitment off the eastern continental shelf of Nova Scotia, Canada, to show that the survival of the larval fish depends on the timing of the local spring bloom of phytoplankton. This link has been suspected for more than 100 years, but its verification has had to wait for technology with sufficient spatial and temporal resolution.

A long-standing hypothesis¹ contends that the abundance of fish year-classes is determined by food availability during the critical period of larval development. Variations in food supply between different years, probably determined by differences between the timing of the spring bloom of phytoplankton and the timing of fish spawning^{2–4}, were thought to account for

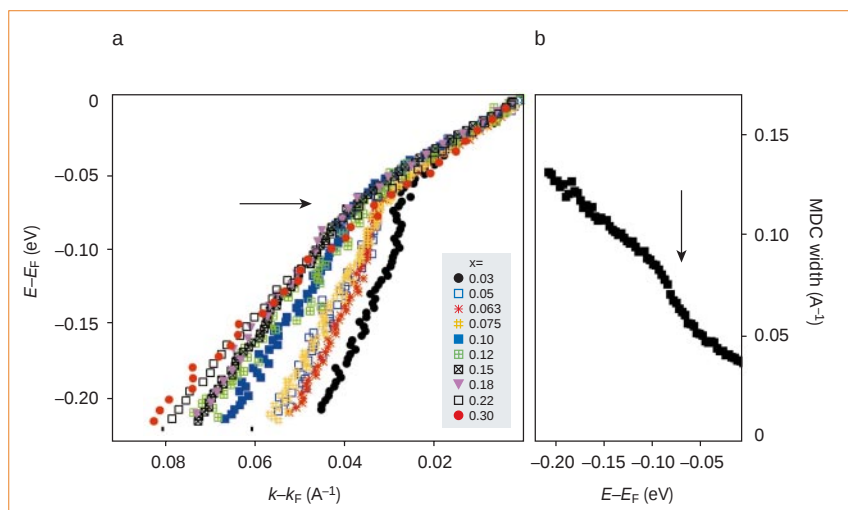


Figure 1 Electron dynamics in the $(\text{La}_{2-x}\text{Sr}_x)\text{CuO}_4$ (LSCO) system. **a**, Dispersion energy, E , as a function of momentum, k , of LSCO with various dopings (where x is between 0.03 (black circles, right curve) and about 0.30 (red circles, left) measured at a temperature of 20 K along the $(0,0) - (\pi,\pi)$ nodal direction. The dispersion is obtained by fitting momentum-distribution curves (MDCs), which represent the photoelectron intensity as a function of momentum, for a given energy. The arrow indicates the position of the kink that separates the dispersion into high-energy and low-energy parts with different slopes. E_F and k_F , Fermi energy and Fermi momentum, respectively. **b**, Scattering rate as measured by MDC width (full width at half-maximum) of the LSCO ($x = 0.063$) sample measured at 20 K. The MDC width is proportionally related to the scattering rate of electrons. The arrow indicates a decrease at an energy of about 70 meV.

later variations in the abundance of adults (the Hjort–Cushing hypothesis).

It has not been feasible to test this possibility at the appropriate scales of time and space by using oceanographic research vessels. Nor was it previously possible to define objective criteria by which an ocean's pelagic ecosystem could be characterized on these scales⁵ before the advent of remote-sensing by satellite of ocean colour, which reveals synoptic fields of phytoplankton biomass at a resolution of about 1 km. Seasonal evolution of phytoplankton fields can be followed with a temporal resolution of 7 days by the construction of composite images. The spring bloom can therefore be characterized with respect to its peak amplitude, timing of peak, timing of initiation and duration.

We focus here on the timing of the bloom peak. A time series showing how this varies between years provides one element for testing the Hjort–Cushing hypothesis. For the other, we need to know how fish stocks varied in the same locality in the same years. From 1970 to the present, haddock (*Melanogrammus aeglefinus*) off the eastern Nova Scotian shelf have been surveyed annually after metamorphosis of the larvae; juveniles (0–2 years old) and adults were captured routinely in the survey. Juvenile abundance, normalized to spawning biomass or total egg production, provides an index of survival and can ultimately reveal the size of the incoming year-class.

For each of the 5,604 pixels (1.5 × 1.5 km each) in the area where eggs and larvae are distributed (Fig. 1a), we determined the timing of the bloom peak in each year. The ocean-colour data cover the periods 1979–81 (CZCS sensor) and 1997–2001 (POLDER sensor for 1997 and SeaWiFS for 1998–2001). Before pooling the two data sets, we expressed each set as a vector of differences from the mean of the set, as a precaution against secular changes between the two data sets. We then tested the null hypothesis that variations in larval survival between years were independent of fluctuations in the timing of the spring phytoplankton bloom. We could account for 89% of the variance in larval survival by variation in the timing of the spring bloom (Fig. 1b). Our analysis is not compromised by differences between sensors or by any systematic error in the chlorophyll-retrieval algorithms for any particular sensor, as we quantified only the timing of events.

Although our ocean-colour time series is short, the data indicate a link between bloom progression and larval survival. The data include two exceptional haddock year-classes (1981 and 1999; Fig. 1b) in the 32-year series covered by the groundfish survey; in both of these years, the spring

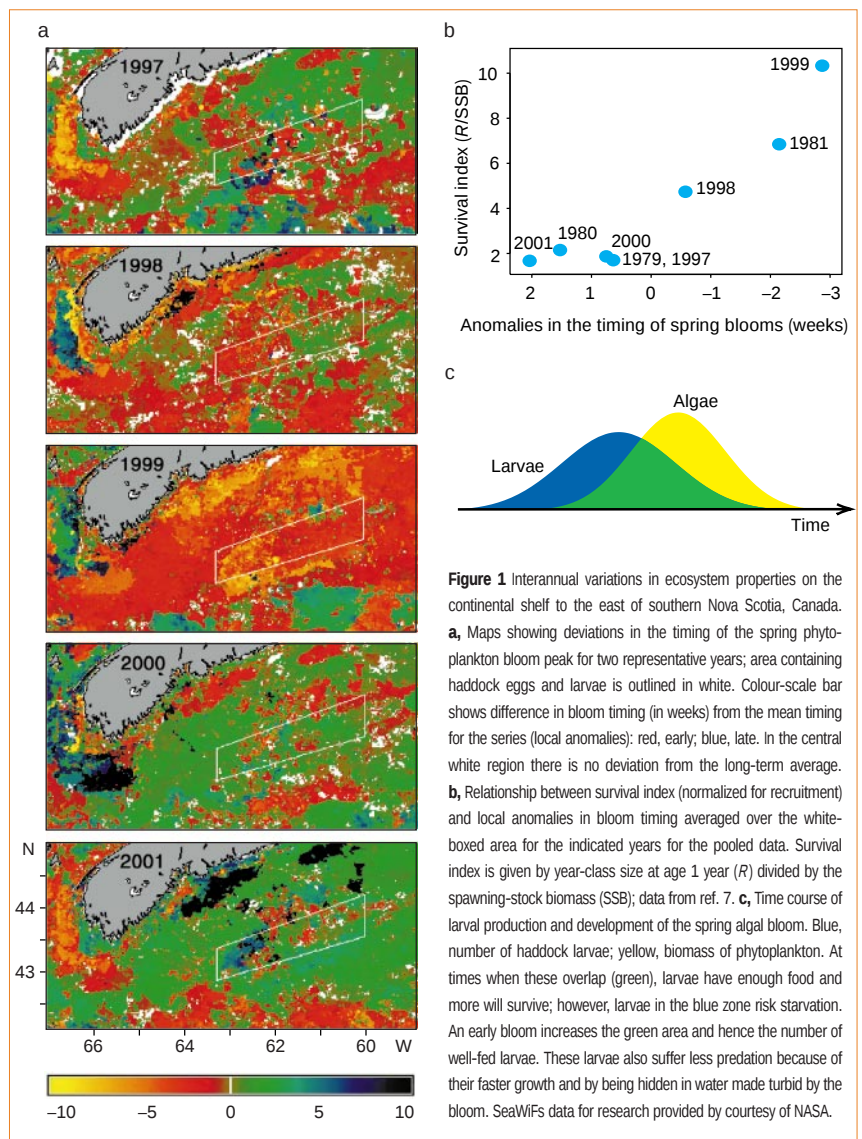


Figure 1 Interannual variations in ecosystem properties on the continental shelf to the east of southern Nova Scotia, Canada. **a**, Maps showing deviations in the timing of the spring phytoplankton bloom peak for two representative years; area containing haddock eggs and larvae is outlined in white. Colour-scale bar shows difference in bloom timing (in weeks) from the mean timing for the series (local anomalies): red, early; blue, late. In the central white region there is no deviation from the long-term average. **b**, Relationship between survival index (normalized for recruitment) and local anomalies in bloom timing averaged over the white-boxed area for the indicated years for the pooled data. Survival index is given by year-class size at age 1 year (R) divided by the spawning-stock biomass (SSB); data from ref. 7. **c**, Time course of larval production and development of the spring algal bloom. Blue, number of haddock larvae; yellow, biomass of phytoplankton. At times when these overlap (green), larvae have enough food and more will survive; however, larvae in the blue zone risk starvation. An early bloom increases the green area and hence the number of well-fed larvae. These larvae also suffer less predation because of their faster growth and by being hidden in water made turbid by the bloom. SeaWiFS data for research provided by courtesy of NASA.

blooms occurred unusually early. Some of the haddock population may therefore have begun to spawn early in the year, even though spawning peaked later. The advantage of an early bloom for a fish species with an extended spawning period might be that fewer of the total larvae produced perish from lack of food (Fig. 1c).

The mean and variance in spawning time presumably respond to environmental pressures in the broadest sense (including predation by humans). Although zooplankton dominate the diet of larval haddock, there is evidence that phytoplankton are also included⁶. Moreover, the Hjort–Cushing hypothesis refers explicitly to phytoplankton, rather than to zooplankton. Although an early spring phytoplankton bloom may not be sufficient to ensure high survival in the same year, it could be a necessary condition.

Our application of remotely sensed ocean-colour data has produced direct

evidence for a putative trophic link that will be an important factor in future analyses of dwindling fish stocks, underlining the need to maintain a continuous, internally consistent stream of such data.

Trevor Platt*, César Fuentes-Yaco*†, Kenneth T. Frank*

**Department of Fisheries and Oceans, Bedford Institute of Oceanography, Ocean Sciences Division, Box 1006, Dartmouth, Nova Scotia B2Y 4A2, Canada*

†Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada e-mail: tplatt@dal.ca

- Hjort, J. *Rapp. Conserv. Explor. Mer* **20**, 1–228 (1914).
 - Cushing, D. H. in *Sea Fisheries Research* (ed. Harden Jones, F. R.) 399–412 (Elek Science, London, 1974).
 - Cushing, D. H. *Adv. Mar. Biol.* **26**, 249–294 (1990).
 - Sverdrup, H. U. J. *Conseil Perm. Int. Explor. Mer* **18**, 287–295 (1953).
 - Platt, T. & Sathyendranath, S. *Int. Coun. Explor. Sea CM O:3* (1996).
 - Kane, J. *Mar. Ecol. Progr. Ser.* **16**, 9–20 (1984).
 - Mohn, R. K. & Simon, J. E. *CSAS Res. Doc.* 2002/102 (2002).
- Competing financial interests:** declared none.

Ecology

Hunting and fox numbers in the United Kingdom

The potential impact of fox-hunting ban in Britain is a contentious issue¹ that has been explored by Baker *et al.*². They conclude that a suspension of lowland fox-hunting for nine months during 2001 made no difference to fox density in certain areas. We are not confident, however, that their analysis supports their conclusions — their study does not consider statistical power or account sufficiently for regional variation, and also uses an inappropriate statistic.

An analysis of statistical power is necessary to demonstrate the probability of detecting an increase in fox density of 10%, 20%, 30%, and so on, given the sample sizes and variability of the data. It seems likely that the power of Baker and colleagues' study to detect plausible effect sizes was low, whether by the analysis published or by any other test: the proper conclusion should be 'no evidence', rather than 'evidence of no effect'.

Moreover, a constant of 1.0 used (twice) in calculating the relative change in faecal abundance, R' , is large by comparison with average scat density, with the result that R' is higher where faecal density before the hunting ban (imposed during the outbreak of foot-and-mouth disease (FMD)) is higher. For the same proportional increase in faecal density, R' takes lower values for a low pre-FMD density than for a high pre-FMD density. For example, a 50% increase from a pre-FMD density of 5 faeces per km² gives $R' = 0.008$, whereas the same 50% increase from a pre-FMD density of 100 per km² gives $R' = 0.085$ (we have used $\log_{10}$; average faecal density was 50 faeces per km² (ref. 3) and the average transect length was 6.9 km (ref. 2)).

If, as hypothesized by Baker *et al.*, pre-FMD fox density had been suppressed in hunted squares, initial faecal density would be low in those squares, but would be predicted to increase during the suspension of hunting. Because of the low initial density, R' would take low values in those squares. In an analysis of all squares, this would tend to mask differences between squares that show an effect and those that show no effect (for no change, $R' = 0$). The consequences of this could influence the conclusions of Baker *et al.*

Finally, an earlier study contrasting three large regions of England and Wales concluded that hunting with dogs was the major part of an effective cull in one region, but not in the other two⁴. It is therefore appropriate to test for regional variation in the impact of hunting pressure, requiring that an interaction term be explicitly modelled. The upland regions where hunting with dogs is more likely to suppress fox densities were not represented in Baker and colleagues' study.

Only when these concerns are addressed

can we concur that Baker *et al.* have clarified the effect of a hunting ban on fox populations. **Nicholas J. Aebischer***, **Sandra E. Baker†**, **Paul J. Johnson†**, **David W. Macdonald†**, **Jonathan C. Reynolds***

*The Game Conservancy Trust, Fordingbridge, Hampshire SP6 1EF, UK

†The Wildlife Conservation Research Unit, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

1. Macdonald, D. W. *et al.* *Managing British Mammals: Case Studies from the Hunting Debate* (Wildlife Conserv. Res. Unit, Oxford, 2000).
2. Baker, P. J., Harris, S. & Webbon, C. C. *Nature* **419**, 34 (2002).
3. The Mammal Society (www.mammal.org.uk).
4. Heydon, M. J. & Reynolds, J. C. *J. Zool.* **251**, 265–276 (2000).

Baker et al. reply — Although there is a tendency for our transformation to result in low R' values for those squares where initial faecal density was low, this does not affect our results. First, this 'bias' is equivalent to a 1.2-fold difference in R' for a 20-fold difference in initial faecal density¹. Second, there was no difference between 'hunted' and 'non-hunted' squares with respect to initial faecal density, direction or magnitude of change in faecal density¹. The starting conditions and pattern of change were therefore the same for hunted and non-hunted squares, and the reservations of Aebischer *et al.* about the transformation are unwarranted.

Our key hypothesis was that hunting with hounds (hereafter termed 'hunting') is additive to other culling practices². Consequently, the absence of hunting during FMD would be expected to result in increased fox abundance in areas that were previously hunted. The most parsimonious way to test this hypothesis, and to remove any possibility of a transformation effect, is by using a signs test³ — this compares the number of squares in which scat density increased with the number in which it decreased, assuming a 50% chance of either event. This negates the need for a regional approach — if hunting is additive, a change of any magnitude in any region would be detected using our paired samples².

To determine the power of this approach, we need the α -error rate ($\alpha = 0.05$), the sample size ($n = 157$; squares with no change are excluded) and the size of the likely effect³. In previous studies^{4–6}, the impact of hunting in Britain (total area, 230,367 km²) has been estimated by extrapolating kill rates per unit area to the total area covered by packs of foxhounds (145,000 km²). The proportion of land in Britain that is hunted is therefore 0.63.

Statistical power is the probability of correctly rejecting a false null hypothesis. For a two-tailed test, the minimum effect size (g) is given by: $0.63 - 0.50 = 0.13$; the statistical power that corresponds to $\alpha = 0.05$, $n = 157$ and $g = 0.13$ is about 0.950 (ref. 3). We therefore had a 95% chance of correctly rejecting a false null hypothesis, for a roughly 13%

deviation from a probability that faecal density would increase (or decrease) in 50% of squares. The actual results showed no change ($P = 0.474$). The observed pattern of variation in faecal density is therefore not consistent with hunting mortality being additive. Furthermore, the small absolute changes in faecal density² indicate minor changes in fox density. We reiterate that these results support the Committee of Inquiry into Hunting with Dogs⁶, which concluded that a permanent ban on hunting is unlikely to result in a dramatic increase in fox numbers.

Philip J. Baker, Stephen Harris, Charlotte C. Webbon

School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK
e-mail: s.harris@bristol.ac.uk

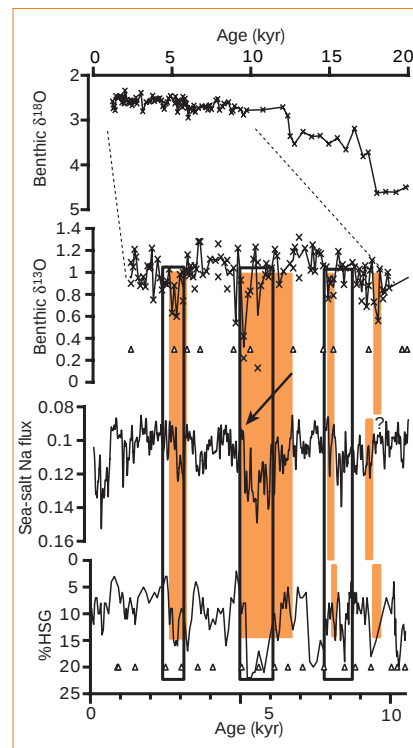
1. The Mammal Society (www.mammal.org.uk).
2. Baker, P. J., Harris, S. & Webbon, C. C. *Nature* **419**, 34 (2002).
3. Cohen, J. *Statistical Power Analysis for the Behavioural Sciences* (Academic, London, 1977).
4. Macdonald, D. W. & Johnson, P. J. in *The Exploitation of Mammal Populations* (eds Taylor, V. J. & Dunstone, N.) 160–207 (Chapman & Hall, London, 1996).
5. Macdonald, D. W. *et al.* *Management and Control of Populations of Foxes, Deer, Hares and Mink in England and Wales, and the Impact of Hunting with Dogs* (Report to Lord Burns' Inquiry into Hunting with Dogs, 2000).
6. Burns, L., Edwards, V., Marsh, J., Soulsby, L. & Winter, M. Report of the Committee of Inquiry into Hunting with Dogs in England and Wales (Stationery Office, London, 2000).

correction

Deepwater variability in the Holocene epoch

D. W. Oppo, J. F. McManus, J. L. Cullen
Nature **422**, 277–278 (2003)

In Fig. 1, the arrow marking the onset and intensification of the 5-kyr event was positioned incorrectly. The correct version of the figure is shown here.



Segregation of object and background motion in the retina

Bence P. Ölveczky^{*†‡}, Stephen A. Baccus[‡] & Markus Meister[‡]

^{*} Division of Health Sciences and Technology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02138, USA

[†] Program in Neuroscience, Harvard University, 220 Longwood Avenue, Boston, Massachusetts 02115, USA

[‡] Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

An important task in vision is to detect objects moving within a stationary scene. During normal viewing this is complicated by the presence of eye movements that continually scan the image across the retina, even during fixation. To detect moving objects, the brain must distinguish local motion within the scene from the global retinal image drift due to fixational eye movements. We have found that this process begins in the retina: a subset of retinal ganglion cells responds to motion in the receptive field centre, but only if the wider surround moves with a different trajectory. This selectivity for differential motion is independent of direction, and can be explained by a model of retinal circuitry that invokes pooling over nonlinear interneurons. The suppression by global image motion is probably mediated by polyaxonal, wide-field amacrine cells with transient responses. We show how a population of ganglion cells selective for differential motion can rapidly flag moving objects, and even segregate multiple moving objects.

Movements of the eye are a fundamental component of vision, as they directly influence the stimulus falling on the retina. There are two main types of eye movement: the large and rapid saccades or pursuit movements by which we redirect our gaze, and the smaller fixational eye movements that occur between saccades^{1,2}. Whereas ballistic gaze-shifting eye movements suppress vision³, small fixational eye movements are essential for seeing: if the retinal image is stabilized, visual perception fades within a tenth of a second⁴.

During fixation, the retinal image drifts over about 0.5° of visual angle, or about 60 cone receptive fields, at an average speed of approximately 0.5 degrees s⁻¹, and any processing of visual information must occur on the background of this drifting motion⁵.

Similar eye movements occur in other animals, including salamander⁶ and rabbit⁷. Despite their importance to vision, the effect of these eye movements on retinal function has received rather limited attention.

Given the presence of continuous eye movements, the fundamental task of detecting object motion within a scene becomes a significant computational problem. The task is not simply to detect motion on the retina; rather, the visual system must discriminate between local motion patterns specific to an object and global motion induced by fixational eye movements⁸. Humans perceive this task as effortless: movements anywhere within a scene immediately ‘pop-out’ and attract our attention⁹, even if their velocity and amplitude is only a fraction of the image motion caused by

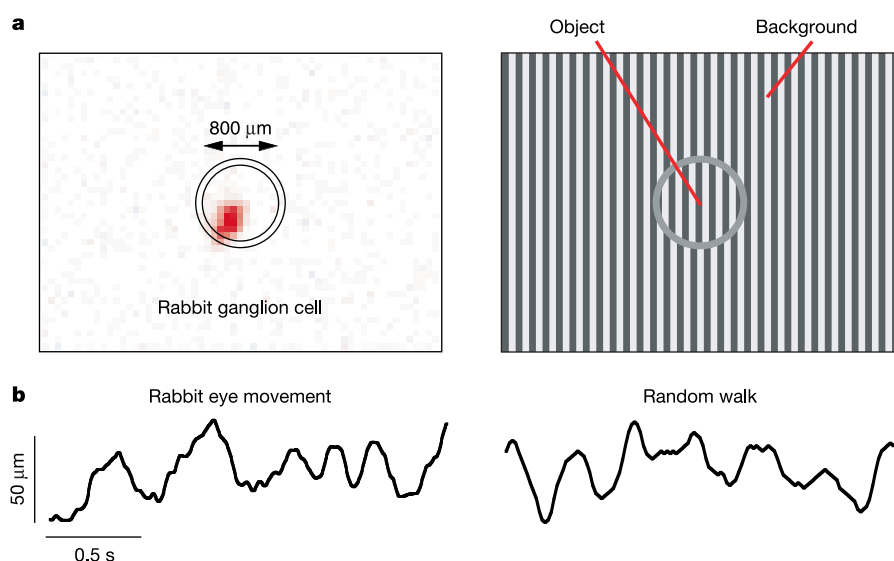


Figure 1 Simulating local object motion on the retina in the presence of fixational eye movements. **a**, Receptive field profile of a rabbit ON Brisk Transient retinal ganglion cell (left; see Methods). A stripe grating representing an object was projected in and around the cell's receptive field centre, while the remainder of the retina was presented with a background grating (right). **b**, Trajectory of vertical fixational eye movements in an unrestrained rabbit, acquired using a scleral search coil⁷ (left). The right panel shows a sample of the random walk trajectory used to jitter the gratings.

fixational eye movements¹⁰. A neuron that subserves this function should respond to local motion on the retina, but only if the motion trajectory differs from that in a large surrounding region. Neurons capable of such computations have been described in the visual cortex^{11,12} and superior colliculus^{13,14} of mammals, and also in the optic tectum of birds¹⁵. However, given that the fixational eye movements in the two eyes differ¹⁶, extraction of object motion probably happens before the visual pathways from the two eyes merge. Here we show that the segregation of object motion and image motion induced by eye movements happens in the retina.

The retina senses differential motion

We recorded the spike trains of ganglion cells in the isolated retina of salamander and rabbit. The stimulus display was divided into a small 'object' region overlying the receptive field centre of the ganglion cell and a surrounding large 'background' region covering the rest of the retina (Fig. 1a). Both object and background were given a visual texture by a simple stripe grating. The background grating jittered laterally with a random walk trajectory, similar to that of fixational eye movements (Fig. 1b). The object grating also jittered in a random walk with the same statistics, either coherently with the background, or incoherently with a different trajectory. The coherent condition simulated the global image motion on the

retina that results when viewing a stationary scene in the presence of eye movements only ('Eye Only' condition). The incoherent condition simulated, in addition, local motion of an object within that scene ('Eye + Object' condition).

In both the salamander and rabbit retina we found ganglion cells that were highly selective for motion within the scene (Fig. 2): these neurons responded vigorously to the Eye + Object condition (Fig. 2a), but were almost completely suppressed under the Eye Only condition (Fig. 2b), even though their receptive field centres experienced the same stimulus under both conditions. When the background region was uniformly grey ('Object Only', Fig. 2c), the responses were similar to the Eye + Object condition (Fig. 2a), indicating that an incoherently moving background has little effect on the centre response. Whereas the stimulus condition in Fig. 2a simulated an object jittering within the scene, a steady drift of the object relative to the background also elicited reliable responses ('Eye + Object Drift', Fig. 2d).

As these ganglion cells are selective for local object motion over global motion, we will refer to them as OMS (object motion sensitive) cells, noting that this class comprises several recognized cell types (for example, 'ON Brisk Transient' cells and 'ON-OFF Direction Selective' cells in rabbit, and 'Fast OFF' cells in salamander; Fig. 2e). Both retinas contain other cell types that show much

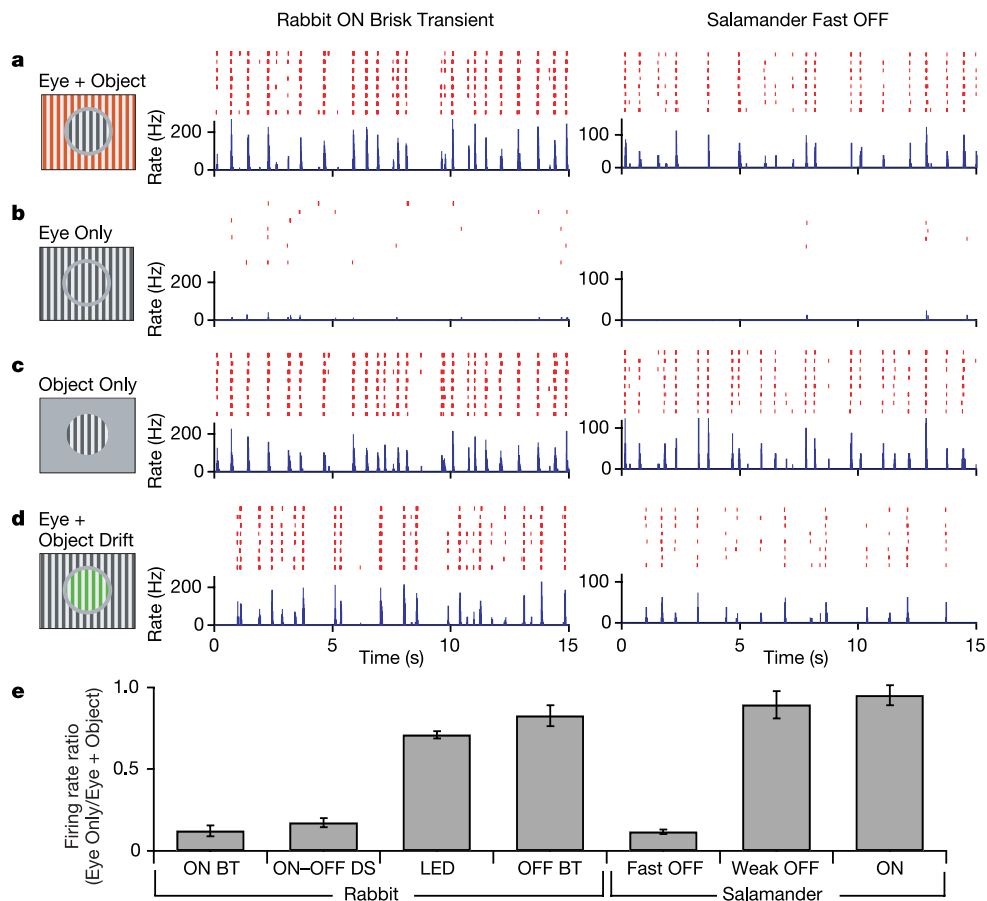


Figure 2 Certain retinal ganglion cells are selective for object motion. **a–d**, Responses to 15 s of jitter from a rabbit ON Brisk Transient cell and a salamander Fast OFF cell. Each panel shows a raster plot with spikes on eight identical stimulus trials (top) and a peri-stimulus time histogram of the firing rate averaged over all trials (bottom). The stimulus conditions are: Eye + Object (**a**), object and background gratings jittered incoherently; Eye Only (**b**), object and background jittered coherently with same trajectory as the object in **a**; Object Only (**c**), object jittered as in **a**, background grey; Eye + Object Drift (**d**),

object and background jittered as in **b** with a steady drift ($450 \mu\text{m s}^{-1}$) added to the object region. **e**, Ratio of firing rates in the coherent (**b**) and incoherent (**a**) motion condition. Data from 6 ON Brisk Transient (ON BT) cells, 11 ON-OFF Direction Selective (ON-OFF DS) cells, 5 Local Edge Detector (LED) cells, and 7 OFF Brisk Transient (OFF BT) cells in two rabbit retinas; 41 Fast OFF cells, 8 Weak OFF cells and 8 ON cells in nine salamander retinas.

smaller, if any, difference between the coherent and incoherent motion conditions (for example, 'OFF Brisk Transient' cells in rabbit and 'ON' cells in salamander, Fig. 2e). Thus, the selectivity for object motion may be a special feature of a few of the parallel pathways that convey retinal output to the brain.

Mechanism of suppression from coherent motion

OMS ganglion cells are excited by motion in or near the receptive field centre, but are suppressed if the same image motion extends over the wider surround. We measured the extent of these antagonistic regions by gradually increasing the size of the object, while keeping the trajectories in the object and background regions different (Fig. 3a). For salamander Fast OFF cells, the firing rate increased up to an object radius of approximately 250 μm ; this is the extent of the region excited by grating motion and corresponds very well to the classic receptive field centre as measured by flashing spots of increasing size (Fig. 3b). As the object grew in size, it began to invade the suppressive surround region, and the response gradually decreased out to radii of about 1,000 μm . Thus, the suppressive effect of coherent motion extends over at least 1 mm on the retina. Applying strychnine largely blocked the antagonistic effect of coherent surround motion, suggesting that it is caused by long-range glycine-mediated inhibition, presumably from wide-field amacrine cells¹⁷.

Ganglion cells can be suppressed by peripheral motion^{18–20}. However, this alone does not explain our findings, as the amount of motion in the background region was identical for both the Eye + Object and the Eye Only conditions, and for all the stimuli in Fig. 3a. As seen in Fig. 2c, the OMS cells respond to a jittering object over the receptive field centre with brief bursts of spikes that are precisely timed to the motion trajectory. We hypothesized that inhibition from peripheral motion arrives in similar brief pulses

that have the same dependence on the motion trajectory as the excitatory events from the centre (Fig. 4a). Under stimulation with coherent motion, the peripheral inhibition would coincide with the excitation from the centre and suppress the cell's response. In response to object motion, when the object and background regions jitter incoherently, inhibition would arrive with random timing

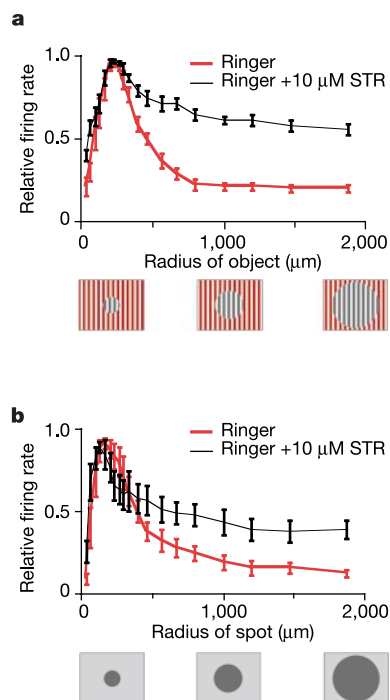


Figure 3 Spatial interactions that produce the sensitivity to object motion. **a**, Relative firing rate of salamander Fast OFF cells ($n = 5$) as a function of object size in the Eye + Object condition. **b**, Relative firing rate of the same cells to a 1-Hz flashing spot of increasing size. The black trace shows the effect of 10 μM strychnine (STR). Firing rates were averaged over 2 min of stimulation and normalized to the maximum firing rate of each cell.

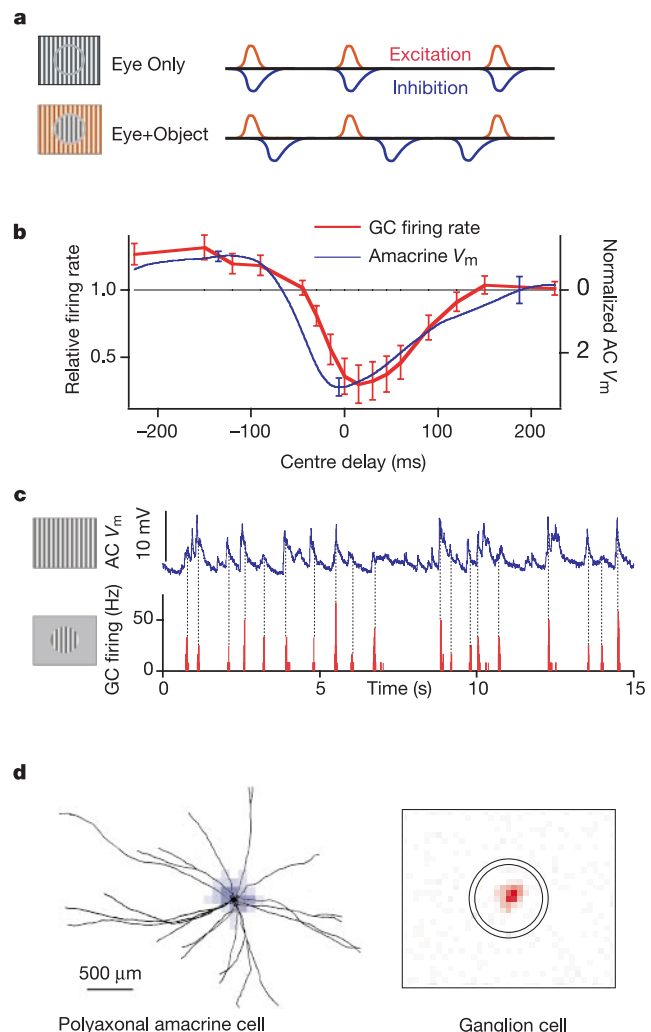


Figure 4 Transient excitation and inhibition are synchronous during coherent motion, causing suppression of firing. **a**, Schematic proposal for the inputs to an OMS ganglion cell: excitation from motion in the receptive field centre, and inhibition from motion in the periphery. Both consist of transient events and are triggered by the same motion features. Under coherent motion they coincide in time, but under incoherent motion they are uncorrelated. **b**, Average firing rate of Fast OFF ganglion cells (GCs) as the jitter trajectories of the object and background regions are shifted in time relative to each other (thick red line). Firing rates were averaged over 5-min stimulus presentations, normalized to the cell's average firing rate under the Object Only condition, then averaged over five neurons. Average membrane potential of polyaxonal amacrine cells (AC V_m) during global (Eye Only) jitter, as a function of time before or after a ganglion cell spike in the Object Only condition using the same trajectory (thin blue line). Each amacrine cell's membrane potential was normalized by subtracting its mean and dividing by its standard deviation, which was 4 ± 1 mV (mean $\pm$ s.e.m.; $n = 3$); note inverted axis, depolarization is downward. **c**, Membrane potential of a polyaxonal amacrine cell in response to coherent motion (top; Eye Only condition). Spiking response of a salamander Fast OFF cell to motion in the Object Only condition, using the same trajectory as for the amacrine cell (bottom). The amacrine and ganglion cells were recorded in different retinas. **d**, Vertical projection of the confocal image of the amacrine cell in **c**, superimposed on its receptive field (left). Receptive field of the ganglion cell in **c**, on the same scale (right).

relative to the excitation and thus be ineffective. To test this idea, we used the same jitter trajectory for the object and background regions, but shifted them with respect to each other in time. As predicted, the suppressive effect was limited to a very brief time window around zero delay (Fig. 4b, thick red trace). This suggests that peripheral inhibition indeed arrives in brief pulses approximately 100 ms wide, triggered by the same motion features as the excitatory events from the centre.

We searched, using intracellular recordings in the salamander retina, for interneurons that might mediate this long-range inhibition. We encountered a type of amacrine cell that responded to coherent jitter (Eye Only) with sharp depolarizations, about 100 ms wide, often carrying action potentials (Fig. 4c). These depolarizing events in amacrine cells aligned perfectly with the excitatory inputs to OMS cells, as marked by the bursts of spikes produced in the Object Only condition (Fig. 4c). If such amacrine cells inhibit the OMS ganglion cells, this could explain the suppression of firing under coherent motion. By calculating the average amacrine cell membrane potential relative to the time of a ganglion cell spike, we predicted how this inhibition should depend on the time delay between object and background motion. Figure 4b shows that the

time course of the amacrine cell membrane potential nicely predicts the measured time course of ganglion cell suppression.

These amacrine cells had visual receptive fields of about 150 μm radius (Fig. 4d), probably mediated by inputs on a small field of dendrites near the soma²¹ (see Supplementary Fig. S1). Several long output processes extended >1 mm across the retina (Fig. 4d). A ganglion cell collecting inhibitory input from these amacrine cell processes will be suppressed whenever the motion in the distant periphery matches the motion over the ganglion cell's receptive field centre. Thus both the anatomy and the physiology of these amacrine cells are consistent with their being the source of inhibition onto OMS cells. Amacrine cells with a similar polyaxonal morphology are found in other species including rabbit^{22,23} and macaque^{24,25}.

Object motion selectivity is independent of spatial pattern

For a neuron to be selective for object motion under natural viewing conditions, the timing of the excitation and inhibition should ideally depend only on the motion trajectory, and be largely independent of the spatial pattern of the stimulus. To test this we changed both the spatial phase and frequency of the object grating, while keeping the background uniformly grey. Indeed, the time

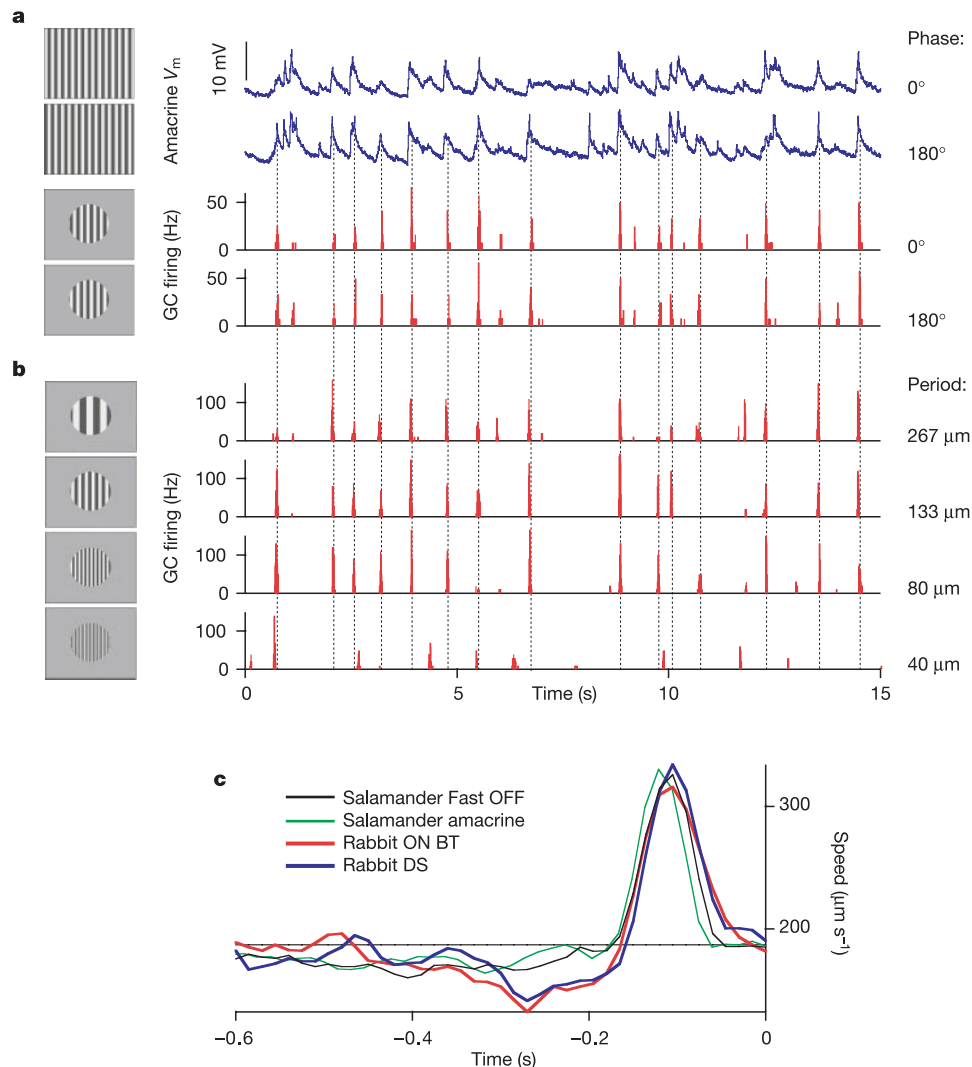


Figure 5 The response of OMS cells is largely independent of the spatial pattern.

a, Responses of a salamander polyaxonal amacrine cell (top) and a Fast OFF ganglion cell (bottom) to a jittering grating (0°) and its contrast-reversed version (180°). **b**, Responses of a different Fast OFF ganglion cell to a grating of varying spatial period. **c**, The average time course of the retinal speed of the jittered grating before a spike for a salamander

polyaxonal amacrine cell (Eye Only condition) and three different types of OMS cells (Object Only condition): the Fast OFF cell in salamander, and the ON Brisk Transient and ON-OFF Direction Selective cells in rabbit. The time-averaged speed for the entire stimulus is marked by the horizontal line.

course of firing of OMS cells did not depend on the spatial phase of the grating: even a 180° phase shift, corresponding to a complete reversal of black and white bars, did not significantly alter the cell's response (Fig. 5a). Nor did such a phase reversal alter the response of polyaxonal amacrine cells (Fig. 5a). Varying the spatial frequency of the grating also had little effect on the firing pattern of the ganglion cells (Fig. 5b), except for very fine gratings with period $<40\text{ }\mu\text{m}$. In that limit, the firing rate declined and the spike bursts were triggered at different time points in the motion trajectory. Polyaxonal amacrine cell responses were similarly robust to changes in spatial frequency (data not shown). These properties differed from those of other amacrine cell types. For example, the responses of sustained OFF-type amacrine cells to coherent motion (Eye Only) were uncorrelated with the excitatory inputs to OMS ganglion cells, and strongly depended on the phase and spatial frequency of the jittered grating (Supplementary Fig. S2).

Whereas the response of OMS cells is largely independent of the spatial pattern, it is almost completely determined by the motion trajectory (Fig. 5a, b). We calculated the average image speed before a spike for three types of OMS cells, and for salamander polyaxonal amacrine cells, in response to a jittered object. The average motion feature that triggered spikes was an acceleration of the grating after a period of slower than average speed (Fig. 5c). For most OMS cells this stimulus was effective regardless of the direction of motion; however, the rabbit ON-OFF Direction Selective cell responded only when the grating accelerated in its preferred direction.

Increasing or decreasing the speed of the jitter by a factor of two did not significantly alter the shape of the preferred speed profile; neither did it alter the OMS cells' sensitivity to differential motion (not shown). This suggests that the function of OMS cells is robust to changes in the statistics of eye movements, which accompany changes in the behavioural state of the animal^{7,26}. Note also that the preferred motion feature for the polyaxonal amacrine cell is

remarkably similar to that which excites the receptive field centres of OMS ganglion cells. This suggests that an inhibitory network involving a single type of amacrine cell could serve to suppress different types of OMS ganglion cells in the same retina.

A model explaining object motion selectivity

The fact that these cells respond to gratings much finer than the receptive field centre, and independently of the phase of the grating, is reminiscent of the Y-type ganglion cells found in cat²⁷ and many other mammals. It is thought that the Y-cell pools excitation from many small subunits in its receptive field, each of which contributes a rectified response²⁸. We implemented a concrete version of this idea (Fig. 6a) and simulated how it would respond to the jittered grating stimulus in the Object Only condition. With just two free parameters, this simple model reproduced with good accuracy the timing of ganglion cell firing in both salamander and rabbit retina (Fig. 6b).

The model's response is independent of the direction of motion and the phase of the grating, because the receptive field centre contains subunits arranged symmetrically in all directions and at all possible phases relative to the grating. Similarly, the response is largely independent of spatial frequency, as long as the bars of the grating are larger than the subunit width. When the stripes become smaller, the model's output declines, and a comparison to the observed responses (Fig. 5b) suggests a minimum subunit width of 20–40 μm . It has been proposed that bipolar cells form the nonlinear subunits of ganglion cell receptive fields^{29,30}. We recorded intracellularly from bipolar cells in the salamander retina. Their receptive field centres measured 35–120 μm ($n = 7$) in width (data not shown), in agreement with the dendritic field size of salamander bipolar cells³¹, and consistent with their role as nonlinear subunits.

The model in Fig. 6a includes an inhibitory network that pools over many nonlinear subunits in the periphery. The fact that the excitatory and inhibitory networks receive input from the same type of subunits results in a selectivity for differential motion between

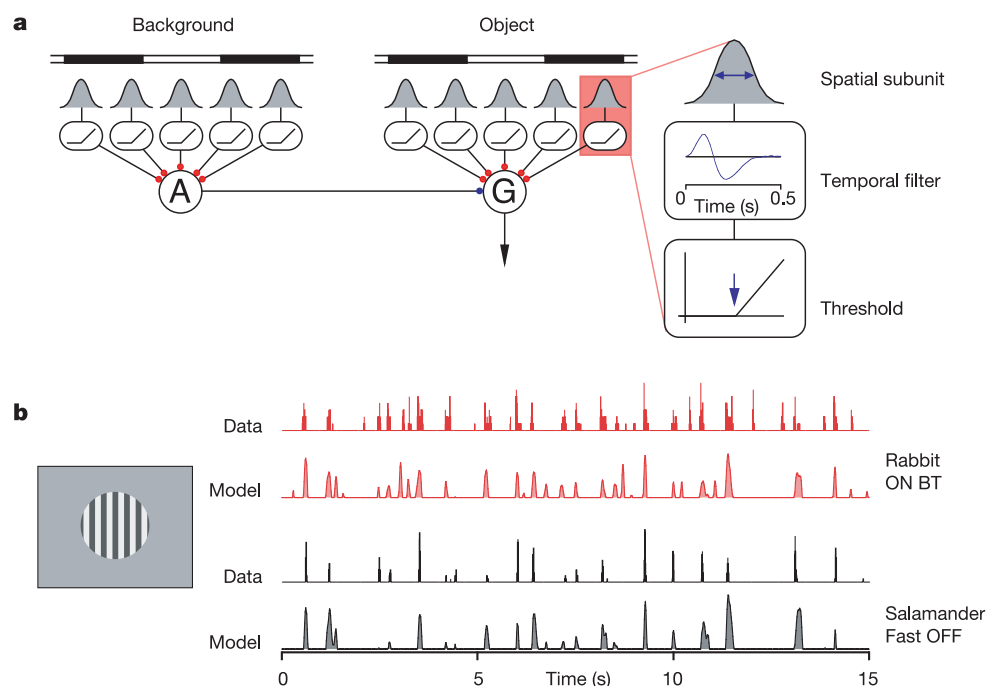


Figure 6 A model of retinal processing that accounts for differential motion sensitivity. **a**, The OMS ganglion cell (G) receives additive excitatory input from many nonlinear subunits underlying the object region. It is also inhibited by amacrine cells (A) that pool over similar nonlinear subunits underlying the background. Each subunit (right) pools light

over a small receptive field, passes the result through a temporal filter, and rectifies the result above a threshold (arrow). **b**, Simulated responses to a jittered grating using the model in **a**, compared to the responses of real OMS cells to the same jitter (see Methods). The stimulus trajectory was the same for both salamander and rabbit cells.

the receptive field centre and the surround. This is accomplished without explicitly computing and comparing the motion vectors in the two regions, contrary to what was found for motion contrast sensors in higher visual centres¹⁵. In fact, the behaviour of this model is not in any way dependent on the direction of motion, only on the speed. To test this prediction, we performed an experiment in which the object and background regions jittered with exactly opposite trajectories, and found that the recorded responses from OMS cells were indeed suppressed, just as under the coherent motion condition (Supplementary Fig. S3). Although this situation represents a departure from the ideal of differential motion detection, it rarely arises in nature, because the relative motion of an object within the scene would have to mimic the observer's eye movements.

Among the rabbit Brisk Transient cells, only the ON-type is suppressed by global motion (Fig. 2e); presumably the OFF-type receives different inhibitory input. For the purpose of detecting object motion it is not essential to have OMS properties in both ON- and OFF-type Brisk Transient cells. Owing to the nonlinear spatial summation, the model of Fig. 6a predicts that an OMS cell would have the same response to a moving pattern whether its subunits are ON-type or OFF-type. This prediction was confirmed experimentally, as the rabbit ON-type OMS cell and the salamander OFF-type OMS cell fired at similar times during the motion trajectory (see Figs 2a–d, 5c and 6b).

Motion pop-out and binding

Finally, we consider how a population of OMS cells represents a visual scene composed of several objects. For this, the stimulus included two objects moving with different trajectories on an incoherently jittering background (Fig. 7a). We recorded the response of many Fast OFF ganglion cells in the salamander retina at more than 200 positions relative to the stimulus display. Figure 7b shows a map of the firing rate in this population. In the region covered by the two moving objects, ganglion cells fired vigorously. In the background region, the cells were suppressed, because they experienced coherent motion between their receptive field centres and the wider periphery. Thus, a population of OMS cells could

support the perceptual 'motion pop-out' effect, which flags local motion within the scene and attracts our visual attention. This pop-out would involve only a short delay: in salamander, the median spike latency of these cells from the onset of object motion was about 230 ms (data not shown).

Sudden movement of an image on the retina is known to synchronize firing in multiple ganglion cells³². Inspection of the spike trains in the two-object experiment showed that OMS cells covering the same object indeed fired in synchrony (Fig. 7c, d), whereas OMS cells seeing differently moving objects were uncorrelated (Fig. 7d). This is expected from the model of Fig. 6a, as the receptive field centres of cells covering the same object experience the same trajectory, and consequently the same speed time course. Because the firing responses are sparse, two neurons belonging to differently moving objects will only rarely fire together by chance. Thus, a group of OMS cells with persistent coincident activity can define a moving object, regardless of its visual pattern or its exact motion trajectory. Segregation of objects based on motion cues is a well-described perceptual phenomenon³³. It has been suggested that synchronous firing is the tag that 'binds' neurons together in an assembly to represent a visual object³⁴. The circuitry of Fig. 6a is a candidate for the underlying neural mechanism.

Discussion

Our experiments involved the retinas of rabbit and salamander, but the essential building blocks required for OMS cells are present in many other species, including primates. Approximately 20% of the magnocellular ganglion cells in the primate retina show nonlinear spatial summation similar to that of our model (Fig. 6a)^{35–37}. Transient, polyaxonal amacrine cells as in Fig. 4d with narrow dendritic and receptive fields and large axonal arborizations have also been found in the macaque retina^{4,25}. Thus, it is probable that ganglion cells with OMS properties exist in many species, including humans. These cells serve as an information channel for object motion and may support diverse functions such as segregating object from background, and directing the gaze towards moving targets.

One would predict that inadvertent stimulation of OMS neurons

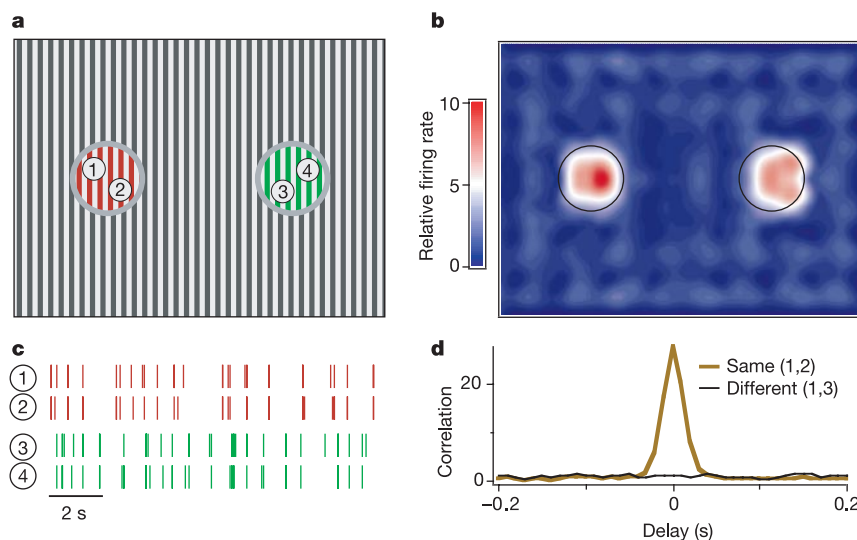


Figure 7 Pop-out of moving objects in a population of OMS ganglion cells. **a**, Schematic of the stimulus: two object regions (red and green, each 800 μm in diameter) are moving with different trajectories on an incoherently jittering background. Encircled numbers denote the receptive field positions of the cells in **c**. **b**, Map of the firing rate in a population of salamander Fast OFF ganglion cells responding to the stimulus in **a** (see Methods).

c, Segment of the spike trains from two pairs of cells covering two differently moving objects: 1 and 2 respond to the object on the left (red), whereas 3 and 4 respond to the object on the right (green). **d**, Cross-correlation function between the spike trains of two cells responding to either the same object (1,2; thick brown line), or differently moving objects (1,3; thin line), normalized by the product of the two cells' mean firing rates.

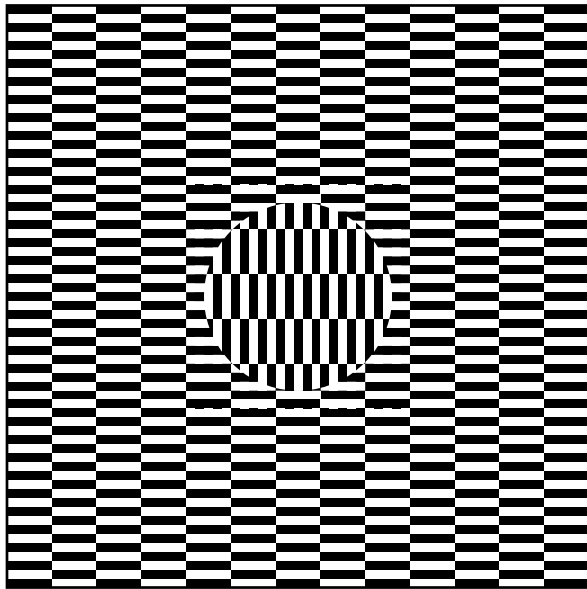


Figure 8 A motion illusion revealed by the Japanese artist Uchi³⁸. The circle appears to float and jitter relative to the background.

should evoke an erroneous perception of object motion. We propose that this happens when viewing certain geometrical patterns, such as the Ouchi illusion³⁸ (Fig. 8). Owing to the unusual geometry of the display, retinal motion of the pattern is mainly governed by vertical eye movements in the periphery, and by horizontal eye movements in the centre. Because the eye executes horizontal and vertical eye movements independently⁵, centre and periphery experience different image motion—analogue to the Eye + Object stimulus (Fig. 2a). This triggers the OMS cells to signal an apparent movement of the circle. Although this is an example of a spatial pattern influencing the response of OMS cells, it should be noted that the geometric peculiarities required for this effect are rare in the natural environment, where retinal function evolved.

We have shown that certain retinal ganglion cells with nonlinear spatial summation have a special role in the presence of fixational eye movements: such a neuron signals when an object in its receptive field centre moves relative to the background, but is almost completely suppressed when the object moves together with the background. This selectivity for object motion in any direction is accomplished by a rather simple mechanism with three crucial ingredients: excitation from many rectified subunits in the receptive field centre; inhibition from the same type of subunits in a wide surround; and the random nature of fixational eye movements that produces transient and sparse activation of both the excitatory and inhibitory networks. □

Methods

Recordings

Retinas of larval tiger salamanders, and Dutch belted and New Zealand white rabbits were isolated in darkness and superfused with oxygenated Ringer's medium at room temperature (salamander) or Ames' solution at 36 °C (rabbit). A piece of retina, 6–8 mm on a side, was placed with the ganglion cell layer facing down on a multi-electrode array, which recorded spike trains simultaneously from many ganglion cells, as described previously³⁹. For intracellular recordings from salamander⁴⁰, sharp microelectrodes were filled with 2 M potassium acetate and 4% neurobiotin, having a final impedance of 150–250 MΩ. Polyaxonal amacrine cell resting membrane potentials ranged from –50 to –75 mV, and their peak responses to jittered gratings were 17 ± 4 mV (mean $\pm$ s.e.m.; $n = 7$) in amplitude. To analyse amacrine cell spiking (Fig. 5c), the action potentials were detected by setting a threshold for the derivative of the membrane potential. After recording, cells were filled iontophoretically (1–5 nA pulses, about 10–15 min), stained with $5 \mu\text{g ml}^{-1}$ streptavidin Alexa-488 (Molecular Probes), and imaged using a confocal microscope with a $\times 40$ oil immersion objective.

Stimulation

Visual stimuli were projected from a computer monitor onto the photoreceptor layer, as described³⁹. All experiments used a mean photopic intensity of 8 mW m^{-2} . Unless otherwise stated, the jittered grating consisted of black and white bars with a periodicity of $133 \mu\text{m}$. The jitter trajectory was generated by stepping the grating randomly in one dimension every 15 ms with a step size of $6.7 \mu\text{m}$. The seeds for generating the random walk in the object and background regions were the same or different for the Eye Only and Eye + Object conditions, respectively. The object region, $800 \mu\text{m}$ in diameter, was separated from the background region, measuring $4,300 \times 3,200 \mu\text{m}$, by a $67\text{-}\mu\text{m}$ grey annulus, except for the experiment in Fig. 3a, where no annulus was present. For Fig. 3b the stimulus was a spot of varying size flashing from black to white at 1 Hz on a grey background. In experiments with different stimulus conditions, the trials were interleaved.

Receptive field mapping

The spatio-temporal receptive fields of all neurons were measured by reverse correlation to a flickering black-and-white checkerboard stimulus³⁹. The spatio-temporal receptive field was approximated as the product of a spatial profile and a temporal filter. The receptive field centre of a ganglion cell was estimated as the region where the spatial profile was larger than one-third of its maximum value. The diameter of the receptive field centre of amacrine and bipolar cells was approximated by the full-width at half-maximum of the two-dimensional gaussian function that best fit the spatial profile.

Cell types

Retinal ganglion cells appear in distinct functional types. For salamander, we classified them on the basis of their spatio-temporal receptive fields⁴¹. We report on the responses of the Fast OFF (~60% of recorded cells), Weak OFF (~12%) and ON cells (~12%). Rabbit ganglion cells were classified on the basis of the spatio-temporal receptive field and the spike train's autocorrelation function, following the criteria of ref. 42. For rabbit, we report on data from ON–OFF Direction Selective cells (~30% of recorded cells), OFF Brisk Transient cells (~20%), ON Brisk Transient cells (~15%) and Local Edge Detectors (~15%). Other cell types in rabbit were encountered rarely, and are not reported.

Analysis

Only ganglion cells with receptive field centres enclosed by the object region were included in the analyses, except for Fig. 3, where only cells with receptive fields concentric with the object region were included. Error bars in figures denote standard error, derived from variation among cells. For calculating the average speed profiles in Fig. 5c, the jitter trajectory was smoothed using a 30-ms box filter. During the experiments for Fig. 7b, the object regions were shifted in increments of $530 \mu\text{m}$ along both the horizontal and vertical dimensions. Responses from 11 salamander Fast OFF cells were analysed, each sampled at 20 independent positions of the stimulus. Each cell's firing rate was normalized with respect to its firing rate under the Eye Only condition. In Fig. 7b, the image value at a given point is the average normalized firing rate of all cells whose receptive field centre contained that point. The results were mirrored on the axis of symmetry in the stimulus, and smoothed using a two-dimensional gaussian filter (standard deviation of $70 \mu\text{m}$).

Simulation

For the simulations in Fig. 6, the subunit width (full width at half maximum of a parabolic profile) was chosen as $42 \mu\text{m}$ for both salamander and rabbit, within the measured range of bipolar cell receptive fields; simulations were robust to changes in this parameter. The waveform of the temporal filter was measured by reverse correlation of ganglion cell spikes to a flicker stimulus⁴³. The threshold was set so that the subunit outputs were non-zero 2.5% (salamander) or 3.5% (rabbit) of the time. The subunits were centred $6.7 \mu\text{m}$ apart, effectively sampling all possible phases relative to the stimulus grating. The region covered by the excitatory subunits was chosen larger than one grating period.

Received 20 December 2002; accepted 18 March 2003; doi:10.1038/nature01652.

Published online 11 May 2003.

1. Yarbus, A. L. *Eye Movements and Vision* (Plenum, New York, 1967).
2. Kowler, E. *Eye Movements and their Role in Visual and Cognitive Processes* (Elsevier, New York, 1990).
3. Ross, J., Morrone, M. C., Goldberg, M. E. & Burr, D. C. Changes in visual perception at the time of saccades. *Trends Neurosci.* **24**, 113–121 (2001).
4. Coppola, D. & Purves, D. The extraordinarily rapid disappearance of entopic images. *Proc. Natl Acad. Sci. USA* **93**, 8001–8004 (1996).
5. Skavenski, A. A., Hansen, R. M., Steinman, R. M. & Winterson, B. J. Quality of retinal image stabilization during small natural and artificial body rotations in man. *Vision Res.* **19**, 675–683 (1979).
6. Manteuffel, G., Plasa, L., Sommer, T. J. & Wess, O. Involuntary eye movements in salamanders. *Naturwissenschaften* **64**, 533–534 (1977).
7. Van der Steen, J. & Collewijn, H. Ocular stability in the horizontal, frontal and sagittal planes in the rabbit. *Exp. Brain Res.* **56**, 263–274 (1984).
8. Gibson, J. J. *The Perception of the Visual World* (Houghton Mifflin, Boston, 1950).
9. Vernon, M. D. *The Psychology of Perception* (Penguin, Baltimore, Maryland, 1962).
10. Graham, C. H. in *Vision and Visual Perception* (ed. Graham, C. H.) 575–588 (Wiley, New York, 1965).
11. Hammond, P. & Smith, A. T. On the sensitivity of complex cells in feline striate cortex to relative motion. *Exp. Brain Res.* **47**, 457–460 (1982).
12. Born, R. T. & Tootell, R. B. Segregation of global and local motion processing in primate middle temporal visual area. *Nature* **357**, 497–499 (1992).
13. Bender, D. B. & Davidson, R. M. Global visual processing in the monkey superior colliculus. *Brain Res.* **381**, 372–375 (1986).
14. Sterling, P. & Wickelgren, B. G. Visual receptive fields in the superior colliculus of the cat. *J. Neurophysiol.* **32**, 1–15 (1969).
15. Frost, B. J. & Nakayama, K. Single visual neurons code opposing motion independent of direction. *Science* **220**, 744–745 (1983).

16. Steinman, R. M. & Collewijn, H. Binocular retinal image motion during active head rotation. *Vision Res.* **20**, 415–429 (1980).
17. Cook, P. B., Lukasiewicz, P. D. & McReynolds, J. S. Action potentials are required for the lateral transmission of glycinergic transient inhibition in the amphibian retina. *J. Neurosci.* **18**, 2301–2308 (1998).
18. Werblin, F. S. Lateral interactions at inner plexiform layer of vertebrate retina: antagonistic responses to change. *Science* **175**, 1008–1010 (1972).
19. Enroth-Cugell, C. & Jakiela, H. G. Suppression of cat retinal ganglion cell responses by moving patterns. *J. Physiol.* **302**, 49–72 (1980).
20. Passaglia, C. L., Enroth-Cugell, C. & Troy, J. B. Effects of remote stimulation on the mean firing rate of cat retinal ganglion cells. *J. Neurosci.* **21**, 5794–5803 (2001).
21. Werblin, F., Maguire, G., Lukasiewicz, P., Eliasof, S. & Wu, S. M. Neural interactions mediating the detection of motion in the retina of the tiger salamander. *Visual Neurosci.* **1**, 317–329 (1988).
22. Famiglietti, E. V. Polyaxonal amacrine cells of rabbit retina: size and distribution of PA1 cells. *J. Comp. Neurol.* **316**, 406–421 (1992).
23. Volgyi, B., Xin, D., Amarillo, Y. & Bloomfield, S. A. Morphology and physiology of the polyaxonal amacrine cells in the rabbit retina. *J. Comp. Neurol.* **440**, 109–125 (2001).
24. Dacey, D. M. Axon-bearing amacrine cells of the macaque monkey retina. *J. Comp. Neurol.* **284**, 275–293 (1989).
25. Stafford, D. K. & Dacey, D. M. Physiology of the A1 amacrine: a spiking, axon-bearing interneuron of the macaque monkey retina. *Vis. Neurosci.* **14**, 507–522 (1997).
26. Grossman, G. E., Leigh, R. J., Bruce, E. N., Huebner, W. P. & Lanska, D. J. Performance of the human vestibuloocular reflex during locomotion. *J. Neurophysiol.* **62**, 264–272 (1989).
27. Hochstein, S. & Shapley, R. M. Linear and nonlinear spatial subunits in Y cat retinal ganglion cells. *J. Physiol.* **262**, 265–284 (1976).
28. Shapley, R. M. & Victor, J. D. Nonlinear spatial summation and the contrast gain control of cat retinal ganglion cells. *J. Physiol.* **290**, 141–161 (1979).
29. Victor, J. D. & Shapley, R. M. The nonlinear pathway of Y ganglion cells in the cat retina. *J. Gen. Physiol.* **74**, 671–689 (1979).
30. Demb, J. B., Zaghloul, K., Haarsma, L. & Sterling, P. Bipolar cells contribute to nonlinear spatial summation in the brisk-transient (Y) ganglion cell in mammalian retina. *J. Neurosci.* **21**, 7447–7454 (2001).
31. Wu, S. M., Gao, F. & Maple, B. R. Functional architecture of synapses in the inner retina: segregation of visual signals by stratification of bipolar cell axon terminals. *J. Neurosci.* **20**, 4462–4470 (2000).
32. Greschner, M., Bongard, M., Rujan, P. & Ammermüller, J. Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nature Neurosci.* **5**, 341–347 (2002).
33. Regan, D. *Human Perception of Objects: Early Visual Processing of Spatial Form Defined by Luminance, Color, Texture, Motion, and Binocular Disparity* (Sinauer, Sunderland, Massachusetts, 2000).
34. Singer, W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24**, 49–65 (1999).
35. Blakemore, C. & Vital-Durand, F. Organization and post-natal development of the monkey's lateral geniculate nucleus. *J. Physiol.* **380**, 453–491 (1986).
36. Kaplan, E. & Shapley, R. M. The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proc. Natl Acad. Sci. USA* **83**, 2755–2757 (1986).
37. Shapley, R., Kaplan, E. & Soodak, R. Spatial summation and contrast sensitivity of X and Y cells in the lateral geniculate nucleus of the macaque. *Nature* **292**, 543–545 (1981).
38. Ouchi, H. *Japanese Optical and Geometrical Art* (Dover, New York, 1977).
39. Meister, M., Pine, J. & Baylor, D. A. Multi-neuronal signals from the retina: acquisition and analysis. *J. Neurosci. Methods* **51**, 95–106 (1994).
40. Baccus, S. A. & Meister, M. Fast and slow contrast adaptation in retinal circuitry. *Neuron* **36**, 909–919 (2002).
41. Warland, D. K., Reinagel, P. & Meister, M. Decoding visual information from a population of retinal ganglion cells. *J. Neurophysiol.* **78**, 2336–2350 (1997).
42. DeVries, S. H. Correlated firing in rabbit retinal ganglion cells. *J. Neurophysiol.* **81**, 908–920 (1999).
43. Chichilnisky, E. J. A simple white noise analysis of neuronal light responses. *Network* **12**, 199–213 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank members of the Meister laboratory for advice; P. Cavanagh, F. Engert, V. Murthy and K. Nakayama for comments on the manuscript; and H. van der Steen for providing the eye movement data in Fig. 1b. This work was supported by a grant from NEI (M.M.) and NRSA (S.A.B.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.M. (meister@fas.harvard.edu).

A role for Wnt signalling in self-renewal of haematopoietic stem cells

Tannishtha Reya^{*†}, Andrew W. Duncan^{*†}, Laurie Ailles[‡], Jos Domen[§], David C. Scherer[‡], Karl Willert^{||}, Lindsay Hintz^{*}, Roel Nusse^{||} & Irving L. Weissman[‡]

^{*} Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina 27710, USA

[‡] Departments of Pathology and Developmental Biology, Stanford University School of Medicine, Stanford, California 94035, USA

[§] Departments of Medicine and Immunology, Duke University Medical Center, Durham, North Carolina 27710, USA

^{||} Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94035, USA

[†] These authors contributed equally to this work

Haematopoietic stem cells (HSCs) have the ability to renew themselves and to give rise to all lineages of the blood; however, the signals that regulate HSC self-renewal remain unclear. Here we show that the Wnt signalling pathway has an important role in this process. Overexpression of activated β -catenin expands the pool of HSCs in long-term cultures by both phenotype and function. Furthermore, HSCs in their normal microenvironment activate a LEF-1/TCF reporter, which indicates that HSCs respond to Wnt signalling *in vivo*. To demonstrate the physiological significance of this pathway for HSC proliferation we show that the ectopic expression of axin or a frizzled ligand-binding domain, inhibitors of the Wnt signalling pathway, leads to inhibition of HSC growth *in vitro* and reduced reconstitution *in vivo*. Furthermore, activation of Wnt signalling in HSCs induces increased expression of *HoxB4* and *Notch1*, genes previously implicated in self-renewal of HSCs. We conclude that the Wnt signalling pathway is critical for normal HSC homeostasis *in vitro* and *in vivo*, and provide insight into a potential molecular hierarchy of regulation of HSC development.

The HSC has the ability to perpetuate itself as well as to differentiate into mature blood cells of all lineages. In the mouse, long-term self-renewing HSCs make up approximately 0.007% of bone marrow and can be isolated by their expression of undetectable levels of lineage markers (such as B220, CD3, Mac-1), high levels of c-Kit and Sca-1, and low levels of Thy-1 (refs 1, 2). Although HSCs have been purified successfully and their phenotypic and functional properties well characterized^{1–4}, a fundamental question that remains is how their self-renewing growth is regulated. On the basis of a screen of genes expressed in HSCs—where we noted that members of the LEF-1/TCF family were expressed (K. Li, S. Cheshier and I.L.W., unpublished data)—and our previous finding that Wnt signalling can influence lymphocyte progenitor cell proliferation⁵, we have investigated whether Wnt signalling influences HSC development.

β -Catenin expression leads to self-renewal of HSCs *in vitro*

We first determined the effects of activating downstream components of the Wnt pathway on HSC function. We activated Wnt signalling in HSCs sorted via fluorescence-activated cell sorting (FACS) (c-Kit⁺ Thy-1.1^{lo} Lin^{−/lo} Sca-1⁺ (KTLS) cells) by retrovirally⁶ transducing them with constitutively active β -catenin⁷. Successful transduction of HSCs with retroviruses requires induction of cell cycle entry through the use of multiple growth factors, which can promote differentiation of stem cells *in vitro*. To minimize the pro-differentiation stimuli encountered by HSCs during infection before experiments of interest, we used HSCs from H2K-BCL-2 transgenic mice, which proliferate to steel factor (SLF) alone⁸. Sorted BCL-2 transgenic HSCs were infected with retroviruses encoding either β -catenin-IRES-GFP (β -catenin, internal ribosome entry site and green fluorescent protein) or IRES-GFP alone, and GFP expression was detected in 45–55% of HSCs, which persisted for the entire *in vitro* culture period (data not shown). GFP-positive (GFP⁺) HSCs were sorted to determine growth kinetics *in vitro* and the ability to reconstitute the immune system *in vivo*.

Short-term growth characteristics of HSCs expressing β -catenin or control vector were determined by cell cycle analysis⁹. In Fig. 1a,

whereas 34% of the HSCs infected with control vector were in S/G2/M phases of the cell cycle, 58% of the HSCs expressing activated β -catenin were in the same phases of the cell cycle (Fig. 1a). To determine whether activated Wnt signalling increased long-term growth, HSCs expressing β -catenin were grown *in vitro* in serum-free medium in the presence or absence of growth factors. Medium containing limiting amounts of SLF allowed the growth of β -catenin-transduced HSCs consistently for at least 8 weeks (Fig. 1b). During this period the GFP⁺ cells underwent eight–nine population doublings to generate at least 100 times the number of input cells. In contrast, HSCs infected with control vector showed minimal growth beyond a two-week period (Fig. 1b). On complete withdrawal of SLF during long-term culture, β -catenin-infected HSCs grew for at least 4 weeks, and in some experiments could be maintained and passaged for as long as 1–2 months. In contrast the control transduced HSCs did not survive beyond 48 h (data not shown).

To determine whether growth in response to activated β -catenin was accompanied by differentiation, the morphological characteristics of these cells were analysed at the end of a two-week period. This time point was chosen to be able to compare the differentiation status of control and β -catenin-transduced HSCs, as the lifespan of HSCs transduced with control vector was limited. Cells infected with control vector were found to have a myelo-monocytic appearance. In contrast, 65–75% of the β -catenin-transduced HSCs had a high nuclear to cytoplasm ratio (Fig. 1c). Consistent with this, although most (75–80%) of the HSCs infected with control vector were positive for lineage markers (Fig. 1d, left panel), only 5–10% of cells infected with β -catenin expressed high levels of lineage markers (predominantly Mac-1, an integrin expressed on fetal HSCs¹⁰ and regenerating HSCs¹¹). In fact, 60% of HSCs infected with β -catenin were lineage-negative and expressed high levels of c-Kit and Sca-1 (Fig. 1d, middle panel) and almost half of these also expressed low levels of Thy-1.1 (right panel). Thus, at least 30% of the cells in β -catenin-transduced cultures had retained the phenotype of HSCs; that is, c-Kit⁺ Thy1.1^{lo} Lin[−] Sca-1⁺ (KTLS cells). This indicated that the expression of activated β -catenin maintained haemato-

poietic stem cells in an immature state, while simultaneously allowing these cells to proliferate, thus expanding the HSC pool 20–48-fold on the basis of the total numbers of cells generated (see Supplementary Information).

To determine whether β -catenin-transduced HSCs retained the functional characteristics of HSCs, and to determine whether there was indeed a functional expansion of HSCs, we tested the ability of β -catenin-transduced HSCs to give rise to sustained production of myeloid, B- and T-cell lineages in lethally irradiated mice when transplanted in limiting numbers^{1,3}. Specifically, HSCs were infected with control or β -catenin virus, cultured for 1 week *in vitro*, transplanted in varying doses down to 125 cells into irradiated mice, and analysed after 11 weeks. As shown in Table 1, 125 β -catenin-transduced HSCs were sufficient to reconstitute the haematopoietic system of all of the irradiated mice. The average level of chimaerism observed was 61% (range of 27.6–86.8%). Furthermore, each of the transplanted mice displayed clear reconstitution along myeloid, T and B lineages (Table 1 and Fig. 2). In contrast, mice receiving 125 HSCs transduced with control vector failed to reconstitute (Table 1). The reconstitution observed at limiting dilution suggested a functional HSC expansion of 8–80-fold in short-term cultures and 96–960-fold in long-term cultures (Supplementary Information). That β -catenin-transduced HSCs lead to higher levels of chimaerism was observed in five independent experiments using different doses of cells and different culture periods (up to 2 months), indicating that β -catenin prevents HSC differentiation while promoting proliferation, thus inducing self-renewal in serum-free medium.

The expansion of HSCs owing to activated β -catenin probably reflects upstream Wnt signals, as we have demonstrated recently that purified Wnt3a causes self-renewal in both BCL-2 transgenic¹² and wild-type HSCs (Supplementary Figs 1 and 2). Specifically, singly plated HSCs generate sixfold or more numbers of progeny in the presence of Wnt3a compared with control conditions¹² (Supplementary Figs 1 and 2). These daughter cells not only maintain an immature phenotype, but also display a 5–50-fold expansion of

HSC function as determined by transplantation analysis of the progeny of single HSCs after expansion *in vitro*¹² (Supplementary Figs 1 and 2).

HSCs *in vivo* normally signal through LEF-1/TCF elements

To determine whether the Wnt signalling pathway is physiologically relevant to HSCs, we tested whether HSCs *in vivo* use signals associated with the Wnt/ β -catenin pathway. HSCs were infected with LEF-1/TCF reporter driving expression of destabilized GFP (TOP-dGFP) or with control reporter with mutated LEF-1/TCF binding sites (FOP-dGFP), and then transplanted into lethally irradiated mice (L. Ailles and I.L.W., manuscript in preparation). Recipient bone marrow was examined after 14 weeks to determine whether donor HSCs demonstrated reporter activity. In the example shown, donor-derived HSCs infected with TOP-dGFP expressed GFP in 28% of the cells (Fig. 3a; range observed 4–28%, mean 11.8%), whereas HSCs from the recipient mouse were negative for GFP (Fig. 3b; range observed 2.3–3.2%, mean 2.7%). Moreover, HSCs transduced with the control reporter did not express GFP significantly, demonstrating that functional LEF-1/TCF binding sites were required for HSC expression of GFP (Fig. 3c). In all cases, no reporter activity was observed in the non-HSC myeloid progenitor fraction (Fig. 3a–c, thin line). As a control, we also tested whether the TOP-dGFP reporter was turned on in response to Wnt3a-mediated signalling in HSCs *in vitro*. Thus, HSCs transduced with either TOP-dGFP or FOP-dGFP were stimulated with Wnt3a, and the extent of GFP expression was monitored. As shown in Fig. 3e, Wnt3a-treated HSCs showed significant reporter activity, demonstrating that the reporter is turned on in response to Wnt stimulus, but not in control conditions. Increased reporter activity was observed when the reporter construct driving non-destabilized GFP was used. These data demonstrate that HSCs in their normal microenvironment respond to endogenous Wnt signalling during self-renewal and/or stimulation into cell cycle, and also support the interpretation that the Wnt3a stimulus that caused increased self-renewal¹² signals through the canonical Wnt pathway.

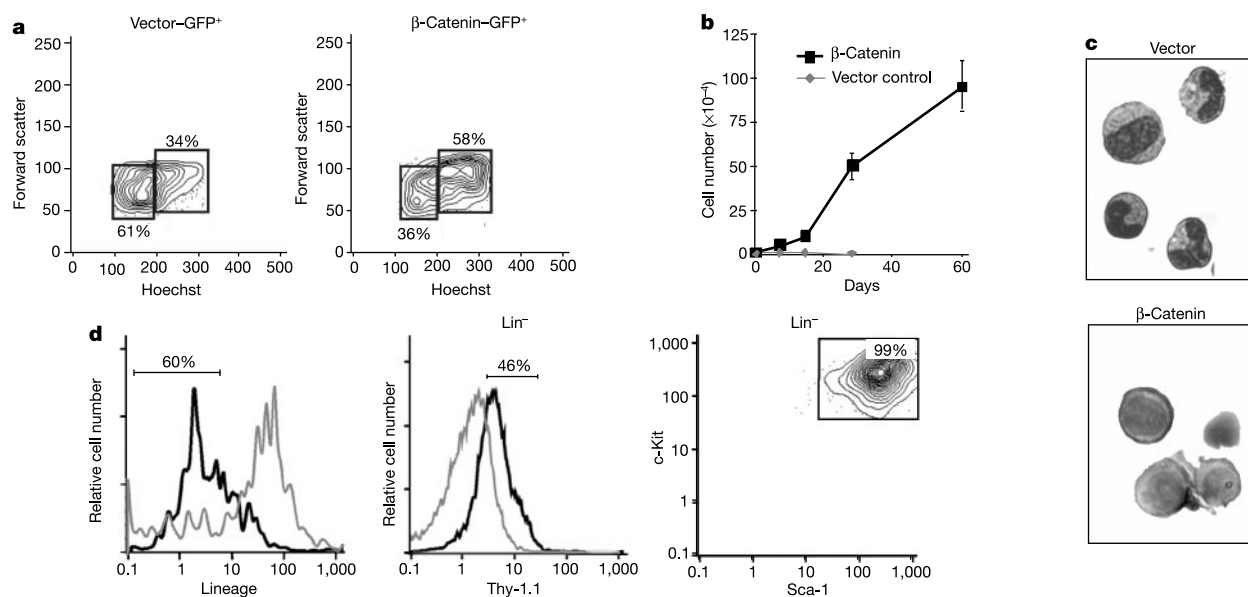


Figure 1 Activated β -catenin promotes growth of HSCs *in vitro* and maintains the immature phenotype of HSCs in long-term cultures. HSCs were infected with activated β -catenin-IRES-GFP or control GFP retrovirus, and subjected to cell cycle analysis after 60 h. **a**, β -Catenin-infected cultures display an increased number of blasting cells (right box, S/G2/M) compared with control. **b**, For long-term growth studies, 10,000 infected HSCs were plated in 1 ng ml^{-1} SLF and monitored over 60 days. Results are from

one of five experiments. **c**, Giemsa staining reveals myeloid characteristics in control cells and HSC morphology (high nucleus to cytoplasm ratio) in β -catenin-infected cells. **d**, Control cells (grey lines) are largely lineage-positive, whereas most β -catenin-infected cells (black lines) are lineage-negative (Lin⁻) or have low levels (left panel). β -Catenin-infected Lin⁻ cells have characteristics of HSCs, including low Thy-1.1 (middle panel), and high c-Kit and Sca-1 (right panel).

Table 1 Reconstitution efficiency of β -catenin- or vector-transduced HSCs

Transplantation		Cell type	Frequency of reconstitution (%)	Chimaerism (%)	
Number	Infection			Average	Range
125	Vector	PB	0 (2/2)	0	0
125	β -Catenin	PB	100 (3/3)	61.2	27.6–86.8
125	β -Catenin	B	100 (3/3)	58.1	37.9–78.8
125	β -Catenin	T	100 (3/3)	19.2	16.0–23.4
125	β -Catenin	M	100 (3/3)	15.0	2.6–25.7

HSCs infected with activated β -catenin or control vector were cultured for a week and cells were injected into groups of three to four lethally irradiated host mice along with 300,000 competing syngeneic bone marrow cells. Cells were isolated from peripheral blood and analysed by flow cytometry after 11 weeks. Average reconstitution rates and ranges of donor cells are shown for peripheral blood (PB), B cell (B), T cell (T) and myeloid (M) lineages. (See also Fig. 2.)

HSCs require intact Wnt signalling

To test whether Wnt signalling is required for normal HSC growth, we used a soluble form of the frizzled cysteine-rich domain (CRD) that inhibits the binding of Wnt proteins to the frizzled receptor^{13,14} (Supplementary Fig. 3). Wild-type HSCs were incubated with growth factors in the presence of IgG–CRD domain fusion protein or control IgG, and cell proliferation was monitored. The presence of the CRD domain inhibited growth of HSCs fourfold compared with control conditions (Fig. 4a). This inhibition provides direct evidence of a Wnt signal modulating HSC survival and proliferation, as soluble CRD acts at the level of Wnt binding the frizzled molecules. Because only HSCs were present, the Wnt signal is probably derived from some or all of the HSCs in the cultures, and is required despite the presence of multiple other growth factors. These results can be interpreted to mean that all HSC mitoses are the result of Wnt signalling, even if the primary signals are not Wnt.

We also inhibited Wnt signalling through an independent inhibitor by ectopically expressing axin in HSCs¹⁵. Axin increases β -catenin degradation and acts as an intracellular inhibitor of Wnt signalling^{16–19}. Live axin-infected wild-type HSCs were re-sorted 48 h after infection and plated in limiting numbers to assay growth in response to a combination of growth factors. Although control-infected cells proliferated 2.3-fold over 60 h, axin-infected cells showed a sevenfold reduction in the total growth response

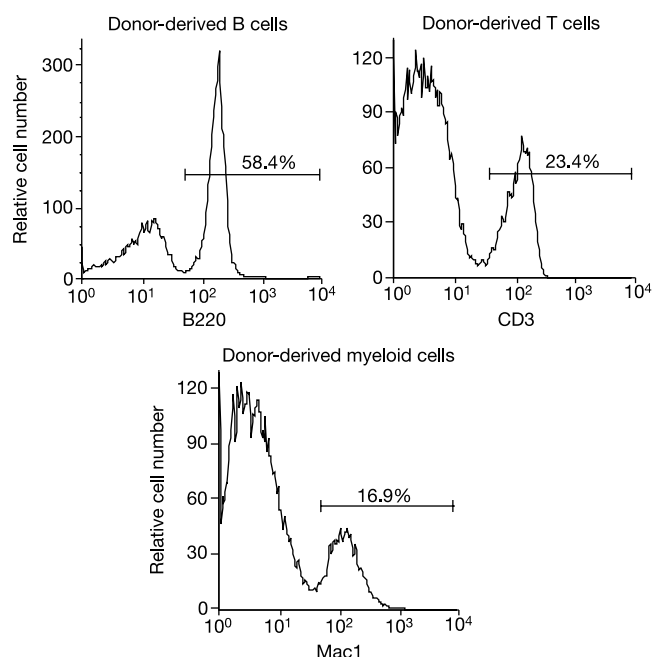


Figure 2 Activated β -catenin-transduced HSCs reconstitute lethally irradiated mice and give rise to multiple haematopoietic lineages. The diagrams show the contribution to each lineage as determined by FACS analysis in a representative sample.

(Fig. 4b). Axin had an inhibitory effect on growth of BCL-2 transgenic HSCs as well (data not shown), which suggests that expression of BCL-2 cannot protect cells from loss of Wnt signalling. To determine whether axin expression had an effect on cell survival, GFP⁺ cells were analysed at the end of the infection period using propidium iodide exclusion. Whereas 80% of the control-infected cells were negative for propidium iodide, only 38% of axin-infected HSCs were negative for propidium iodide (Fig. 4c), indicating that axin expression has significant effect on cell survival by blocking β -catenin function.

To determine whether Wnt signalling is required for haematopoietic stem cell responses *in vivo*, we injected axin- or control-transduced viable HSCs into lethally irradiated mice and analysed the level of reconstitution after 10 weeks. Mice transplanted with control-infected HSCs displayed on average sevenfold greater chimaerism (reconstitution range 5–11.6%) than mice transplanted with axin-infected HSCs (reconstitution range 0–1.8%) (Fig. 4e). A representative example of contribution from axin- or vector-infected HSCs in transplanted mice is shown in Fig. 4d. These data show that inhibition of the Wnt pathway reduces reconstitution, suggesting that Wnt signalling is required for normal development of HSCs *in vivo*. This finding, together with the finding

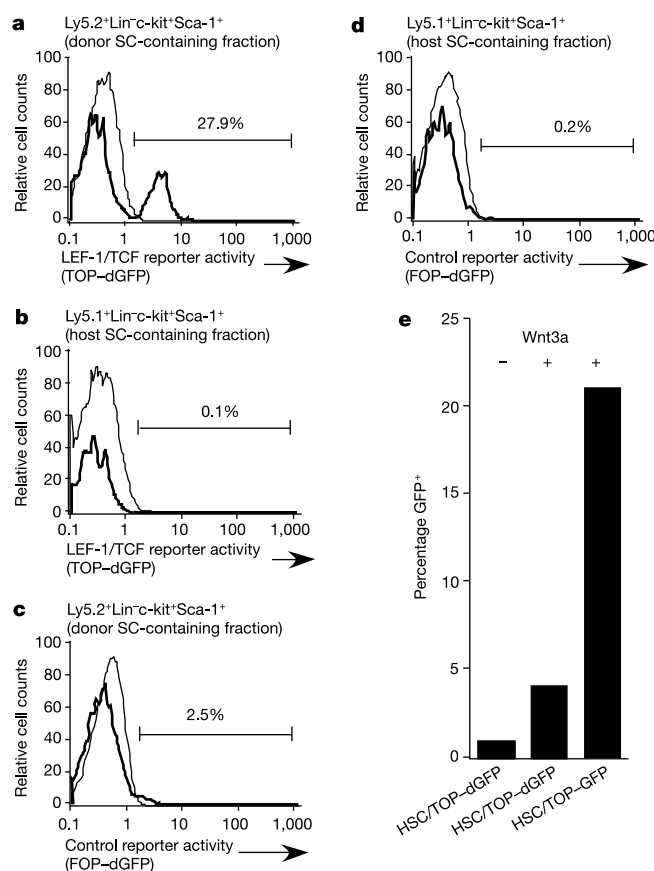


Figure 3 HSCs respond to Wnt signalling in native bone marrow microenvironment. HSCs were infected with a lentiviral reporter containing either LEF-1/TCF binding sites linked to destabilized GFP (TOP-dGFP), or mutated LEF-1/TCF binding sites linked to destabilized GFP (FOP-dGFP). Infected HSCs were transplanted into three lethally irradiated recipient mice, and analysed after 14 weeks. The data shown represent two independent experiments. **a, b**, GFP expression is shown in donor-derived (**a**) or host-derived (**b**) HSCs. **c, d**, Donor-derived HSCs carrying mutated LEF-1/TCF reporter (**c**) as well as the recipient mouse HSCs (**d**) are GFP negative. Expression of GFP in donor-derived Lin[−] c-Kit⁺ Sca-1[−] cells (non-HSCs) is shown by thin lines (**a–d**). **e**, HSCs infected with TOP-dGFP or TOP-GFP (a non-destabilized GFP) were stimulated *in vitro* with control medium or with 100 ng ml^{−1} Wnt3a, and the extent of GFP expression measured.

that HSCs respond to Wnt signalling *in vivo* (Fig. 3), indicates that Wnt/ β -catenin signalling is an important physiological mediator of HSC-derived haematopoiesis.

β -catenin upregulates HoxB4 and Notch1 in HSCs

We wished to determine whether Wnt signalling might be regulating HSC self-renewal by upregulating genes previously implicated in HSC self-renewal. To this end we tested upregulation of HoxB4 and Notch1 (refs 20, 21). By using real-time polymerase chain reaction (PCR) analysis on HSCs infected with either β -catenin or control vector, we found that HoxB4 was upregulated an average of 3.5-fold and Notch1 was upregulated 2.5-fold (Fig. 5a). In contrast, Gapdh expression was not differentially regulated as a consequence of β -catenin expression, and was used as a control (Fig. 5b). These data show that genes so far identified as regulators of HSC self-renewal may be related and perhaps act in a molecular hierarchy.

Discussion

Our study shows that components of the Wnt signalling pathway

can induce proliferation of purified KTLS bone marrow HSCs while significantly inhibiting their differentiation, thereby resulting in functional self-renewal. We find that expression of β -catenin in HSCs results in increased growth with significantly reduced differentiation *in vitro* for a period of at least many weeks. HSCs transduced with β -catenin give rise to sustained reconstitution of myeloid and lymphoid lineages *in vivo*, when transplanted in limiting numbers. We also find that Wnt signalling is required for the growth response of normal HSCs to other cytokines, as over-expression of axin leads to reduced stem cell growth both *in vitro* and *in vivo*. Furthermore, the inhibition of HSC growth with frizzled-CRD and the finding that Wnt3a causes expansion of HSCs supports the interpretation that the effects of β -catenin and axin reflect upstream Wnt activity. Finally, studies with HSCs containing a LEF-1/TCF reporter indicate that HSCs *in vivo* respond to endogenous Wnt stimulation. The expression of a number of Wnt proteins in the bone marrow⁵ and frizzled receptors in bone-marrow-derived progenitors and HSCs supports this possibility²².

Most growth factors that act on HSCs in culture induce no or limited expansion²³ or are unable to prevent differentiation^{8,24}. Thus, one of the most notable findings of our work is the induction of proliferation and the prevention of HSC differentiation by the Wnt signalling pathway. Other signals that increase proliferation of HSCs include Notch²⁰ and sonic hedgehog²⁵. Moreover, the cyclin-dependent kinase inhibitor p21^{Cip1/Waf1} (ref. 26) and the transcription factor HoxB4 (ref. 21) have been shown to be involved in regulating self-renewal of HSCs. Notably, Wnt signalling has been shown to interact with many of these pathways in a variety of organisms^{27–30}, and our data show that both HoxB4 and Notch1 are upregulated in response to Wnt signalling in HSCs. This raises the possibility that the effects of Wnt signalling on HSCs are mediated through HoxB4 and/or Notch1. Whereas HoxB4 could act directly on these HSCs as demonstrated previously²¹, Notch1 action would require nearby Notch1 ligands.

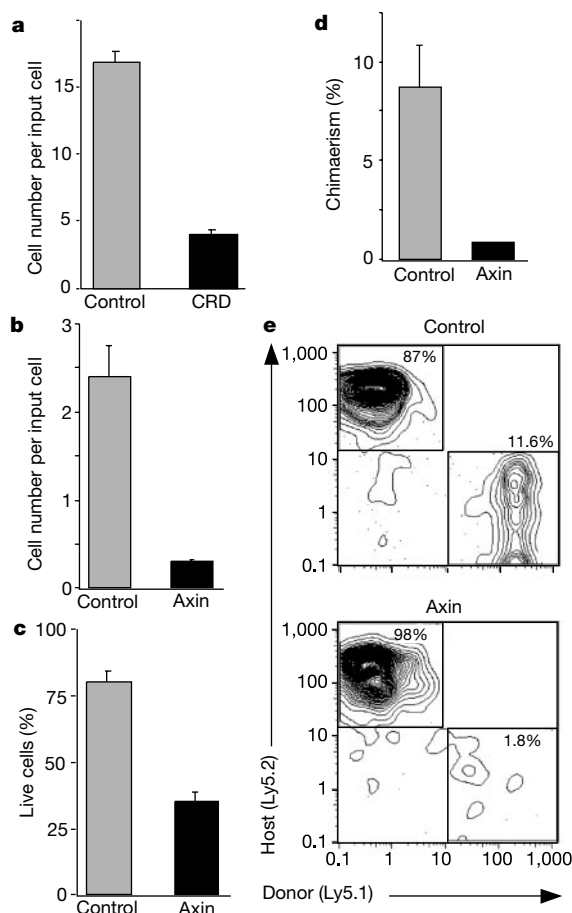


Figure 4 Inhibition of Wnt signalling reduces growth of HSCs *in vitro* and inhibits reconstitution *in vivo*. **a**, HSCs (20 cells per well) were cultured for 60 h in medium containing mitogenic factors and either IgG-CRD or control IgG. **b**, HSCs were infected with virus encoding axin-IRES-GFP or GFP alone. Growth of infected HSCs in the presence of mitogenic factors was monitored over 60 h. **c**, The number of live cells was determined by propidium iodide staining. **d**, **e**, The development of HSCs *in vivo* was determined by injecting 1,000 control or axin-infected cells per mouse into groups of four lethally irradiated, allelically marked (Ly5.2) host mice along with 300,000 competing syngeneic bone marrow cells. Cells were isolated from peripheral blood and analysed by flow cytometry after >10 weeks. Donor-derived (Ly5.1⁺) cells were monitored in the peripheral blood of hosts; analysis from a representative recipient and average reconstitution is shown.

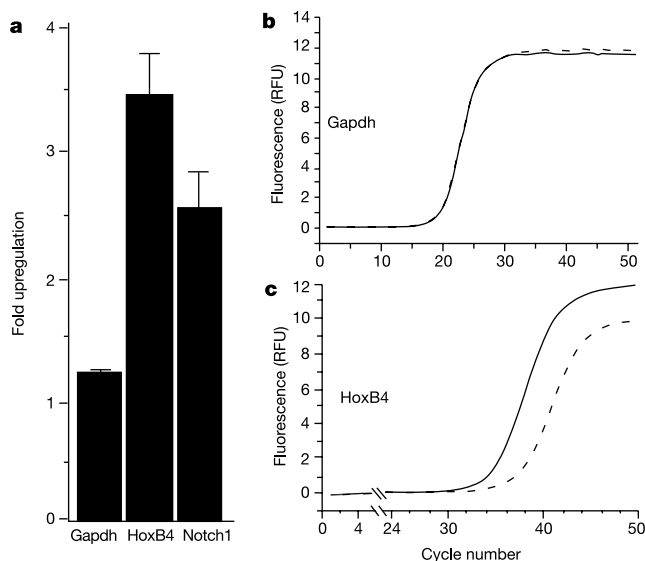


Figure 5 HSCs expressing β -catenin upregulate HoxB4 and Notch1. **a**, Purified wild-type HSCs were infected with activated β -catenin-IRES-GFP or control vector-IRES-GFP, and infected cells sorted based on GFP expression at 48 h. The RNA isolated from these cells was reverse transcribed and expression of HoxB4 and Notch1 was analysed by real-time PCR analysis. Results are averaged over five independent PCR reactions. **b**, **c**, Representative graphs of real-time PCR analysis demonstrating equal amounts of Gapdh (**b**) and differential amounts of HoxB4 (**c**) products from β -catenin-transduced HSCs (solid line) and control-transduced HSCs (dashed line). RFU, relative fluorescence units.

The ability of Wnt3a to induce expansion of HSCs is consistent with previous studies showing increased numbers of haematopoietic progenitors from mouse fetal liver and human bone marrow cells stimulated with Wnt-containing supernatants^{31,32}. These studies, although consistent with ours, are more difficult to interpret as they did not use purified HSCs, lacked *in vivo* reconstitution analysis, and did not provide evidence of a physiological requirement for Wnt signalling for HSCs. Components of the Wnt pathway have also been shown to promote proliferation of primitive cells in the skin^{33,34}, the gut^{35,36} and the brain³⁷, and to inhibit differentiation to a variety of lineages in embryonic stem cells³⁸, raising the possibility that Wnt signalling may be used as a general cue for self-renewal in stem and/or progenitor cells from diverse tissues. Its role as a self-renewal signal does not, however, preclude its involvement in differentiation of stem cells in certain contexts³⁹.

Our findings may have important implications for human haematopoietic cell transplantation. We have found that soluble Wnt3a protein induces proliferation of highly purified human bone marrow HSCs in the absence of any other growth factor (T.R., T. Miyamoto and I.L.W., unpublished observation). Induction of HSC growth by Wnt signalling may allow *in vitro* expansion of a patient's own or an allogeneic donor's HSCs, and could provide an increased source of cells for future transplantation. Finally, we have previously raised the hypothesis that self-renewal is a property that could be dangerous, as an adequate definition of cancer stem cells is poorly regulated self-renewal of a particular stage of a developmental lineage⁴⁰. The demonstration here that the Wnt/ β -catenin pathway may have a role in haematopoietic stem cell self-renewal leads us to propose that this pathway should be studied for a role in self-renewal of cancer stem cells. □

Methods

Mice

C57BL/Ka Ly5.1, Thy-1.1 (wild-type and BCL-2), C57BL/Ka Ly5.2, Thy-1.1, and AKR/J mice were used at 6–10 weeks of age. Mice were bred and maintained on acidified water in the animal care facility at Stanford and Duke University Medical Centers.

HSC isolation

We sorted HSCs from mouse bone marrow as described⁴¹. All cell sorting and FACS analysis was carried out on a FACS Vantage (Becton Dickinson) at the Stanford shared FACS facility and the Duke Cancer Center FACS facility. Cells were sorted and reanalysed on the basis of expression of c-Kit, Sca-1, low levels of Thy-1.1, and low to negative levels of lineage markers (Lin).

Cell cycle analysis

Retrovirally transduced HSCs were collected from cultures and stained with Hoechst 3342 (Molecular Probes) at 37 °C for 45 min in Hoechst medium⁹. Cells were then washed and analysed by Flow cytometry to determine the cell cycle profile of GFP⁺ cells.

Viral production and infection

Virus was produced by triple transfection of 293T cells with murine stem cell virus constructs along with gag-pol and vesicular stomatitis virus G glycoprotein constructs. Viral supernatant was collected for three days and concentrated 100-fold by ultracentrifugation at 50,000g. For viral infection, 10,000 HSCs were sorted into wells of a 96-well plate and cultured overnight in the presence of SLF (30 ng ml⁻¹) for BCL-2 transgenic HSCs (used in Fig. 1, Fig. 2 and Table 1), or SLF (30 ng ml⁻¹) plus TPO (30 ng ml⁻¹) for wild-type HSCs (all other figures and Supplementary figures). After 12 h, concentrated retroviral supernatant was added to the cells at a 1:1 ratio. Cells were then incubated at 32 °C for 12 h and 37 °C for 36 h before GFP⁺ cells were sorted for *in vitro* and *in vivo* assays. Lentiviruses used were produced as previously described. Briefly, 293T cells were transfected with the transfer vector plasmid, the VSV-G envelope-encoding plasmid pMD.G, and the packaging plasmid CMVΔR8.74 (ref. 42). The supernatant was collected and concentrated by ultracentrifugation. All cytokines were purchased from R&D systems.

In vitro HSC proliferation assays

Freshly purified or virally transduced HSCs were plated at 1 to 20 cells per well in Terasaki plates. Cells were sorted into wells containing serum-free medium (X-vivo15, BioWhittaker) supplemented with 5 × 10⁻⁵ M 2-mercaptoethanol and the indicated growth factors. Proliferation was monitored by counting the number of cells in each well at defined intervals. For longer-term cultures, transduced HSCs were plated in 96-well plates in the absence or presence of SLF (1 ng ml⁻¹), and the number of cells generated was monitored by cell counting at defined intervals. For inhibition of growth by CRD or axin, cells were cultured in the presence of mitogenic factors (SLF (30 ng ml⁻¹), Flt-3L (30 ng ml⁻¹), interleukin-6 (10 ng ml⁻¹)).

In vivo analysis of HSC function

Virally transduced HSCs were cultured *in vitro* and injected retro-orbitally into groups of 4–6 congenic recipient mice irradiated with 9.5 Gy using a 200-kV X-ray machine, along with 300,000 rescuing host total bone marrow or Sca-1-depleted bone marrow cells. Host mice were given antibiotic water after irradiation. Transplanted mice were bled at regular periods to determine the percentage of the haematopoietic compartment contributed by donor cells. Donor and host cells were distinguished by allelic expression of CD45 (Ly5) or expression of the BCL-2 transgene.

Lentiviral reporter assays

The enhanced GFP (eGFP) or the d2-eGFP gene (destabilized, half-life of 2 h; Clontech) was cloned downstream of a LEF-1/TCF-responsive promoter, containing three LEF-1/TCF binding motifs and a TATA box⁴³. This cassette was then cloned into a self-inactivating lentiviral vector plasmid, and virus was produced as described above.

For *in vivo* assays, HSCs were transduced with reporter lentiviruses and cultured in X-Vivo15 with glutamate, 5 × 10⁻⁵ M 2-mercaptoethanol, and a cocktail of cytokines including 10 ng ml⁻¹ interleukin-11, 10 ng ml⁻¹ TPO, 50 ng ml⁻¹ SCF, 50 ng ml⁻¹ Flt-3L. Cells were incubated at 37 °C for 6 h overnight and transplanted into lethally irradiated congenic recipients. Lethally irradiated mice received 500 transduced HSCs along with rescue bone marrow. For analysis, haematopoietic progenitor cells were analysed for reporter activation 14–24 weeks after transplantation.

For *in vitro* assays, purified HSCs were sorted directly into medium (IMDM/10% FBS plus interleukin-11, TPO, SCF and Flt-3L, as above) and plated at 500–1,000 cells per well in 96-well plates. Individual wells were transduced with the appropriate lentiviral reporter and stimulated with or without purified Wnt3a (about 100 ng ml⁻¹). Cells were collected 5 days later, stained with propidium iodide to exclude non-viable cells, and analysed for GFP expression.

Real-time PCR analysis

A total of 75,000 HSCs cultured in 96-well plates containing X-Vivo15, 5 × 10⁻⁵ M 2-mercaptoethanol and 100 ng ml⁻¹ SLF were infected with either β -catenin or control lentiviruses. After two days in culture, transduced cells were isolated on the basis of GFP expression. RNA was prepared using Trizol (Invitrogen) and linearly amplified using a modified Eberwine synthesis⁴⁴. Each amplified RNA was converted to the first strand and analysed for differential gene expression by real-time PCR. Complementary DNAs were mixed with FastStart Master SYBR Green polymerase mix (Roche), primers (Supplementary Information) and real-time PCR was performed using a LightCycler (Roche).

Received 12 December 2002; accepted 27 March 2003; doi:10.1038/nature01593.

Published online 27 April 2003.

- Spangrude, G. J., Heimfeld, S. & Weissman, I. L. Purification and characterization of mouse hematopoietic stem cells. *Science* **241**, 58–62 (1989); erratum **244**, 1030 (1989).
- Uchida, N. & Weissman, I. L. Searching for hematopoietic stem cells: evidence that Thy-1.1^{lo} Lin⁻ Sca-1⁺ cells are the only stem cells in C57BL/Ka-Thy-1.1 bone marrow. *J. Exp. Med.* **175**, 175–184 (1992).
- Morrison, S. J. & Weissman, I. L. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity* **1**, 661–673 (1994).
- Weissman, I. L. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* **100**, 157–168 (2000).
- Reya, T. *et al.* Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity* **13**, 15–24 (2000).
- Hawley, R. G., Fong, A. Z., Burns, B. F. & Hawley, T. S. Transplantable myeloproliferative disease induced in mice by an interleukin 6 retrovirus. *J. Exp. Med.* **176**, 1149–1163 (1992).
- Barth, A. L., Stewart, D. B. & Nelson, W. J. T cell factor-activated transcription is not sufficient to induce anchorage-independent growth of epithelial cells expressing mutant β -catenin. *Proc. Natl Acad. Sci. USA* **96**, 4947–4952 (1999).
- Domen, J. & Weissman, I. L. Hematopoietic stem cells need two signals to prevent apoptosis; BCL-2 can provide one of these, Kit/c-Kit signaling the other. *J. Exp. Med.* **192**, 1707–1718 (2000).
- Cheshier, S. H., Morrison, S. J., Liao, X. & Weissman, I. L. *In vivo* proliferation and cell cycle kinetics of long-term hematopoietic stem cells. *Proc. Natl Acad. Sci. USA* **96**, 3120–3125 (1999).
- Morrison, S. J., Hemmati, H. D., Wandycz, A. M. & Weissman, I. L. The purification and characterization of fetal liver hematopoietic cells. *Proc. Natl Acad. Sci. USA* **92**, 10302–10306 (1995).
- Randall, T. D. & Weissman, I. L. Phenotypic and functional changes induced at the clonal level in hematopoietic stem cells after 5-fluorouracil treatment. *Blood* **89**, 3596–3606 (1997).
- Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* advance online publication, 27 April 2003 (doi:10.1038/nature01611).
- Hsieh, J. C., Rattner, A., Smallwood, P. M. & Nathans, J. Biochemical characterization of Wnt-frizzled interactions using a soluble, biologically active vertebrate Wnt protein. *Proc. Natl Acad. Sci. USA* **96**, 3546–3551 (1999).
- Cadigan, K. M., Fish, M. P., Rulifson, E. J. & Nusse, R. Wingless repression of *Drosophila* frizzled 2 expression shapes the Wingless morphogen gradient in the wing. *Cell* **93**, 767–777 (1998).
- Zeng, L. *et al.* The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell* **90**, 181–192 (1997).
- Sakanaka, C., Weiss, J. B. & Williams, L. T. Bridging of β -catenin and glycogen synthase kinase-3 β by axin and inhibition of β -catenin-mediated transcription. *Proc. Natl Acad. Sci. USA* **95**, 3020–3023 (1998).
- Ikedo, S. *et al.* Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3 β and β -catenin and promotes GSK-3 β -dependent phosphorylation of β -catenin. *EMBO J.* **17**, 1371–1384 (1998).
- Yamamoto, H. *et al.* Axil, a member of the Axin family, interacts with both glycogen synthase kinase 3 β and β -catenin and inhibits axis formation of *Xenopus* embryos. *Mol. Cell Biol.* **18**, 2867–2875 (1998).
- Hart, M. J., de los Santos, R., Albert, I. N., Rubinfeld, B. & Polakis, P. Downregulation of β -catenin by

- human Axin and its association with the APC tumour suppressor, β -catenin and GSK3 β . *Curr. Biol.* **8**, 573–581 (1998).
20. Varnum-Finney, B. *et al.* Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling. *Nature Med.* **6**, 1278–1281 (2000).
21. Antonchuk, J., Sauvageau, G. & Humphries, R. K. HOXB4-induced expansion of adult hematopoietic stem cells *ex vivo*. *Cell* **109**, 39–45 (2002).
22. Ivanova, N. B. *et al.* A stem cell molecular signature. *Science* **298**, 601–604 (2002).
23. Miller, C. L. & Eaves, C. J. Expansion *in vitro* of adult murine hematopoietic stem cells with transplantable lympho-myeloid reconstituting ability. *Proc. Natl Acad. Sci. USA* **94**, 13648–13653 (1997).
24. Audet, J., Zandstra, P. W., Eaves, C. J. & Piret, J. M. Advances in hematopoietic stem cell culture. *Curr. Opin. Biotechnol.* **9**, 146–151 (1998).
25. Bhardwaj, G. *et al.* Sonic hedgehog induces the proliferation of primitive human hematopoietic cells via BMP regulation. *Nature Immunol.* **2**, 172–180 (2001).
26. Cheng, T. *et al.* Hematopoietic stem cell quiescence maintained by p21cip1/waf1. *Science* **287**, 1804–1808 (2000).
27. Hing, H. K., Sun, X. & Artavanis-Tsakonas, S. Modulation of wingless signaling by Notch in *Drosophila*. *Mech. Dev.* **47**, 261–268 (1994).
28. Maloof, J. N., Whangbo, J., Harris, J. M., Jongeward, G. D. & Kenyon, C. A Wnt signaling pathway controls hox gene expression and neuroblast migration in *C. elegans*. *Development* **126**, 37–49 (1999).
29. Hooper, J. E. Distinct pathways for autocrine and paracrine Wingless signalling in *Drosophila* embryos. *Nature* **372**, 461–464 (1994).
30. Ingham, P. W. & Hidalgo, A. Regulation of wingless transcription in the *Drosophila* embryo. *Development* **117**, 283–291 (1993).
31. Austin, T. W., Solar, G. P., Ziegler, F. C., Liem, L. & Matthews, W. A role for the Wnt gene family in hematopoiesis: expansion of multilineage progenitor cells. *Blood* **89**, 3624–3635 (1997).
32. Van Den Berg, D. J., Sharma, A. K., Bruno, E. & Hoffman, R. Role of members of the Wnt gene family in human hematopoiesis. *Blood* **92**, 3189–3202 (1998).
33. Zhu, A. J. & Watt, F. M. β -catenin signalling modulates proliferative potential of human epidermal keratinocytes independently of intercellular adhesion. *Development* **126**, 2285–2298 (1999).
34. Gat, U., DasGupta, R., Degenstein, L. & Fuchs, E. *De novo* hair follicle morphogenesis and hair tumors in mice expressing a truncated β -catenin in skin. *Cell* **95**, 605–614 (1998).
35. Korinek, V. *et al.* Depletion of epithelia stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature Genet.* **19**, 1–5 (1998).
36. van de Wetering, M. *et al.* The β -catenin/Tcf-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002).
37. Chenn, A. & Walsh, C. A. Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science* **297**, 365–369 (2002).
38. Kielman, M. F. *et al.* Apc modulates embryonic stem-cell differentiation by controlling the dosage of β -catenin signaling. *Nature Genet.* **32**, 594–605 (2002).
39. Huelsken, J., Vogel, R., Erdmann, B., Cotsarelis, G. & Birchmeier, W. β -Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin. *Cell* **105**, 533–545 (2001).
40. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
41. Domen, J., Cheshier, S. H. & Weissman, I. L. The role of apoptosis in the regulation of hematopoietic stem cells: Overexpression of BCL-2 increases both their number and repopulation potential. *J. Exp. Med.* **191**, 253–364 (2000).
42. Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* **72**, 8463–8471 (1998).
43. Korinek, V. *et al.* Constitutive transcriptional activation by a β -catenin-Tcf complex in APC^{-/-} colon carcinoma. *Science* **275**, 1784–1787 (1997).
44. Wang, E., Miller, L. D., Ohnmacht, G. A., Liu, E. T. & Marincola, F. M. High-fidelity mRNA amplification for gene profiling. *Nature Biotechnol.* **18**, 457–459 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank W. J. Nelson and A. Barth for providing the β -catenin construct; G. Nolan and L. Naldini for providing the retroviral and lentiviral packaging systems; F. Costantini for providing the axin construct; J. Nathans for the Fzd8-CRD IgG construct; H. Clevers and B. Vogelstein for providing the modified LEF-1/TCF elements used in the reporter constructs; A. Carlton for help with proliferation assays; and P. Feliciano for generation of retroviral constructs. We are grateful to L. Jerabek and P. Shahi for laboratory management and technical help; L. Hidalgo for animal care; S. Smith and V. Braunstein for preparation of antibodies; M. Cook for cell sorting; and M. Krangel for use of the Light Cycler. D.C.S. is a fellow of the American Cancer Society, L.A. is a fellow of the Canadian Institute of Health Research and R.N. is an investigator of the Howard Hughes Medical Institute. This work was supported by NIH grants awarded to I.L.W., and funds from the Cancer Research Institute, Leukemia Research Foundation and the NIH awarded to T.R.

Competing interests statement The authors declare competing financial interests: details accompany the paper on www.nature.com/nature.

Correspondence and requests for materials should be addressed to T.R. (t.reya@duke.edu).

Polarization of the prompt γ -ray emission from the γ -ray burst of 6 December 2002

Wayne Coburn* & Steven E. Boggs*†

* Space Sciences Laboratory, and † Department of Physics, University of California, Berkeley, California 94720, USA

Observations of the afterglows of γ -ray bursts (GRBs) have revealed that they lie at cosmological distances, and so correspond to the release of an enormous amount of energy^{1,2}. The nature of the central engine that powers these events and the prompt γ -ray emission mechanism itself remain enigmatic because, once a relativistic fireball is created, the physics of the afterglow is insensitive to the nature of the progenitor. Here we report the discovery of linear polarization in the prompt γ -ray emission from GRB021206, which indicates that it is synchrotron emission from relativistic electrons in a strong magnetic field. The polarization is at the theoretical maximum, which requires a uniform, large-scale magnetic field over the γ -ray emission region. A large-scale magnetic field constrains possible progenitors to those either having or producing organized fields. We suggest that the large magnetic energy densities in the progenitor environment (comparable to the kinetic energy densities of the fireball), combined with the large-scale structure of the field, indicate that magnetic fields drive the GRB explosion.

We used the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI)³ to make these γ -ray observations of GRB021206. RHESSI has an array of nine large-volume (300 cm³) coaxial germanium detectors with high spectral resolution, designed to study solar X-ray and γ -ray emission (3 keV–17 MeV). RHESSI has high angular resolution (2'') in the $\sim 1^\circ$ field of view of its optics; however, the focal plane detectors are unshielded, open to the whole sky. Thus, while the chances are small that RHESSI will see a GRB in its imaging field of view, it measures them frequently in the focal plane detectors themselves. These observations provide high-resolution spectra, individual photon times and energies, as well as the potential for polarization measurements. RHESSI is not optimized to act as a γ -ray polarimeter, but several aspects of its design make it the most sensitive instrument to date for measuring astrophysical γ -ray polarization.

In the soft γ -ray range of ~ 0.15 –2.0 MeV, the dominant photon interaction in the RHESSI detectors is Compton scattering. A small fraction of incident photons will undergo a single scatter in one detector before being scattered and/or photoabsorbed in a second separate detector, which are the events sensitive to the incident γ -ray polarization. Linearly polarized γ -rays preferentially scatter in directions perpendicular to their polarization vector. In RHESSI, this scattering property can be used to measure the intrinsic polarization of astrophysical sources. The sensitivity of an instrument to polarization is determined by its effective area to scatter events, and the average value of the polarimetric modulation factor^{4,5}, $\mu(\theta, E_\gamma)$, which is the maximum variation in azimuthal scattering probability for polarized photons. This factor is given by $\mu = (d\sigma_\perp - d\sigma_\parallel)/(d\sigma_\perp + d\sigma_\parallel)$, where $d\sigma_\perp$, $d\sigma_\parallel$ are the Klein–Nishina differential cross-sections for Compton scattering perpendicular and parallel to the polarization direction, respectively, which is a function of the incident photon energy E_γ , and the Compton scatter angle θ between the incident photon direction and the scattered photon direction. For a source of count rate S and fractional polarization Π_s , the expected azimuthal scatter angle distribution is $dS/d\phi = (S/2\pi)[1 - \mu_m \Pi_s \cos(2(\phi - \eta))]$, where ϕ

is the azimuthal scatter angle, η is the direction of the polarization vector, and μ_m is the average value of the polarimetric modulation factor for the instrument. Although RHESSI has a small effective area (~ 20 cm²) for events that scatter between detectors, it has a relatively large modulation factor in the 0.15–2.0 MeV range, $\mu_m \approx 0.2$, as determined by Monte Carlo simulations described below.

In comparison with other γ -ray instruments (COMPTEL, BATSE) that have attempted to measure polarization in the past^{5,6}, RHESSI has the major advantage of quickly rotating around its focal axis (centred on the Sun) with a 4-s period. Rotation averages out the effects of asymmetries in the detectors and passive materials that could be mistaken for a modulation. Because polarimetric modulations repeat every 180°, any source lasting more than half a rotation (2 s) will be relatively insensitive to the systematic uncertainties that typically plague polarization measurements. Finally, although the RHESSI detectors have no positioning information themselves, they are relatively loosely grouped on the spacecraft, allowing the azimuthal angle for a given scatter to be determined to within $\Delta\phi = 13^\circ$ r.m.s. This angular uncertainty will decrease potential modulations by a factor of 0.95, which is included in our calculated modulation factor.

Prompt γ -ray emission from GRB021206 was detected with RHESSI on 6 December 2002 at 22:49 UT (Fig. 1). This GRB was also observed⁷ with the Interplanetary Network (IPN), which reported a 25–100 keV fluence of 1.6×10^{-4} erg cm⁻², and a peak flux of 2.9×10^{-5} erg cm⁻² s⁻¹, making this an extremely bright GRB. The IPN localized⁸ GRB021206 to a 57 square-arcminute

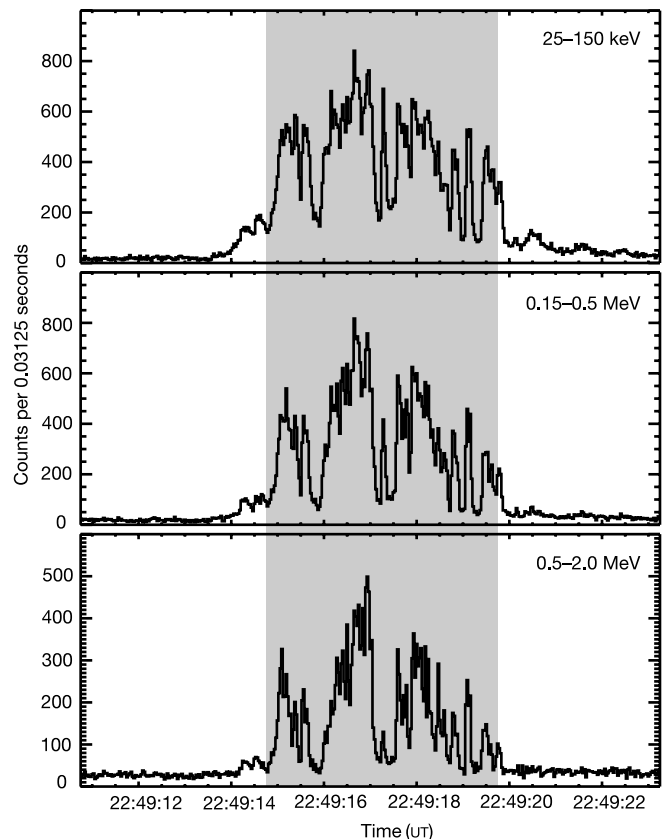


Figure 1 RHESSI light curves (in total measured counts) in three energy bins for GRB021206. The IPN localized⁸ this GRB to 18° off solar, which precluded optical afterglow searches; however, the brightness, duration, and proximity to the RHESSI rotation axis made it an ideal candidate to search for polarization. The shaded region shows our 5-s integration time for the polarization analysis.

error box located 18° from the Sun. For RHESSI, we analysed photons in the energy range 0.15–2.0 MeV that scattered between two, and only two, detectors for the 5-s integration period shown in Fig. 1. Scattered photons constitute roughly 10% of the total 0.15–2.0 MeV light-curve events. Counts were binned by the centre-to-centre azimuthal angle between the two detectors around the RHESSI roll axis, corrected for the rotation of the spacecraft at the time of the photon event. This azimuthal distribution is plotted in Fig. 2. The top panel shows the raw data, as well as the expected variation for an unpolarized GRB due only to the light-curve variability. The bottom panel shows the residual of the data once this unpolarized distribution is subtracted. The residual shows a large modulation, which we interpret as a linear polarization of $\Pi_m = 80 \pm 20\%$. This observation is the first astrophysical polarization measurement at γ -ray energies. Note that the uncertainty on this polarization amplitude reflects in large part our uncertainty in the modulation factor. The fact that we have measured a polarization somewhere in this range has a significance of $<10^{-8}$ (or a confidence $>5.7\sigma$).

Linear polarization is generally considered a clear indication of synchrotron emission. For electrons with an energy spectrum characterized by a power-law distribution with spectral index p , synchrotron photons are emitted with a linear polarization⁹ of $\Pi = (p+1)/(p+7/3)$. For shock acceleration¹⁰, typical values of $p = 2$ –3 correspond to linear polarizations of 70–75%. For unresolved sources, polarizations from many directions generally add together to produce net polarizations that are a fraction of this maximum value. While the source is unresolved in our observation, if the source electrons are moving with a bulk Doppler factor of Γ , we are only viewing the source over a solid angle $\Omega_\gamma \approx 1/\Gamma^2$. Thus, for typical values of $\Gamma \geq 300$ that have been implied from GRB afterglow observations, we are effectively resolving a source region of solid angle $\Omega_\gamma \approx 10^{-5}$ sr. Our measurement of this high polarization is consistent with synchrotron origin for the initial GRB from a region of nearly uniform magnetic field. For this emission process to be radiatively efficient as is implied by many afterglow observations^{11,12}, the magnetic energy densities must be close to equipartition (comparable to the kinetic energy densities of the fireball)^{13,14}.

This polarization from the prompt γ -ray emission is significantly higher than optical polarizations of 1–3% typically measured from

afterglows^{15–17}, as well as the optical polarization of 10% recently reported¹⁸ from the afterglow of GRB020405. The afterglow emission implies a strong magnetic field behind the shocks¹⁹, although the field energy density can be well below equipartition²⁰. However, the implied magnetic field strengths are too large to have been due to a progenitor field being dragged along by the expanding fireball²¹, or compression of the ISM magnetic fields in external shocks²². Therefore, the magnetic fields responsible for the afterglow synchrotron emission are probably turbulent fields that have built up behind the shocks^{21,23}, which is consistent with the relatively small optical afterglow polarizations. This locally generated field invoked for the afterglow has influenced many researchers to consider the magnetic field responsible for the prompt burst of γ -rays as a turbulent, fireball-induced field as well. In the standard internal shock model²⁴, the prompt γ -rays are produced by synchrotron emission of electrons accelerated to relativistic energies by shock acceleration, requiring magnetic energy densities near equipartition in the progenitor environment. However, late-time turbulent fields have no direct implications on whether the fields responsible for the prompt γ -ray emission are predominantly turbulent or organized^{21,25}. Our observation conclusively shows that the engine driving the GRB has a strong, large-scale magnetic field.

Another potential source of polarization would be as follows: if unpolarized γ -rays are initially beamed into a small-angle jet, and then scatter at an angle θ into our line of sight. The polarization induced by this Compton scattering⁹ is $\Pi = (1 - \cos^2 \theta)/(1 + \cos^2 \theta)$ for photon energies below ~ 0.1 MeV, decreasing as the photon energy approaches and exceeds the electron rest mass, 0.511 MeV. For our energy band, we estimate a maximum Compton-induced polarization of 70% for scatter angles near 90° . However, this process would be inefficient, requiring a much larger γ -ray energy budget than the synchrotron case. First, the distribution of scattered γ -rays would be nearly isotropic, requiring an energy budget $4\pi/\Omega_j$ larger than if the observed γ -rays originated from a collimated jet of opening solid angle Ω_j . Second, to maintain such a high polarization as we observed, the γ -rays must undergo only a single scatter into our direction because secondary scatters will erase the induced polarization from the initial scatter. This condition requires that the scatter medium be optically thin, $\tau \ll 1$, and that the luminosity of the unseen initial γ -ray beam be larger than the observed scattered γ -ray emission by a factor $\sim 1/\tau$. Because Compton-scattering-induced polarization requires a total γ -ray luminosity several orders of magnitude larger than that implied by synchrotron emission, and elaborate source geometries, the synchrotron origin for the polarization is preferred.

We suggest that our observation is evidence that the magnetic fields are actually powering the GRB explosion itself. It has been argued that a ‘passive’ magnetic field—that is, a field dragged from the surface of central object with a magnetic dipole moment, but not driving the GRB—could not be strong enough to produce the prompt γ -ray emission without an additional locally generated turbulent magnetic field²⁵. If this conclusion holds, our observation is consistent with models of a magnetically driven GRB fireball, where the driving magnetic field was generated by extracting the rotational energy of an accretion disk around a central compact object through differential rotation^{26,27}, by directly extracting the spin energy of a black hole threaded by magnetic field lines²⁸, or by extracting the spin energy of a highly magnetized neutron star²⁹. Alternatively, our observation of a large-scale magnetic field could support models of dynamos in the post-shock flows³⁰ if the shocks can be shown to be unstable on large size scales and on timescales comparable to the prompt γ -ray emission. $\square$

Methods

The top panel of Fig. 2 shows the azimuthal scatter distribution of 0.15–2.0 MeV events for the 5-s integration period shown in Fig. 1, corrected for the rotation of the spacecraft at the time of each photon event. This distribution is the sum of three components: the GRB

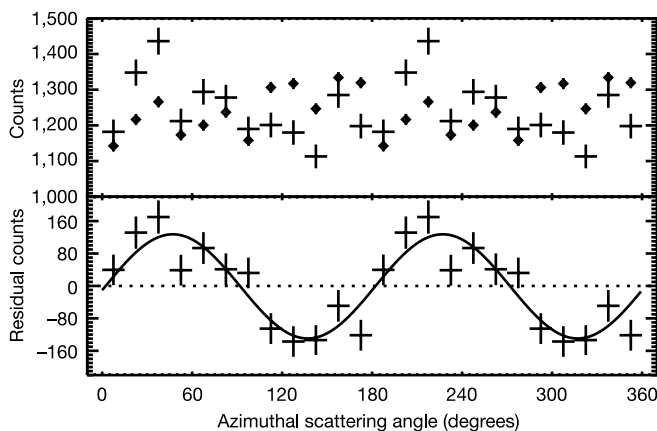


Figure 2 The azimuthal scatter distribution for the RHESSI data, corrected for spacecraft rotation. Counts were binned in 15° angular bins between 0° – 180° , and plotted here twice for clarity. The top plot shows the raw measured distribution (crosses), as well as the simulated distribution for an unpolarized source (diamonds) as modelled with a Monte Carlo code, given the time-dependent incident flux. The bottom plot shows the RHESSI data with the simulated distribution subtracted. This residual is inconsistent with an unpolarized source (dashed line) at a confidence level $>5.7\sigma$. The solid line is the best-fit modulation curve, corresponding to a linear polarization of $80 \pm 20\%$.

scatter event rate which averages 820 ± 8 counts per bin, the chance coincidence rate which averages 374 ± 6 counts per bin, and the background scatter event rate of 49 ± 2 counts per bin. These rates were determined and confirmed using event rates before and after the burst, combined with studies of the readout times of multiple detectors during scattered events, and were verified independently using our Monte Carlo simulations. Although the rotation will average out systematic variations in the scatter angle distribution, we still have to correct for the complex time profile of the burst itself, which will cause variations for an unpolarized source owing to the finite number of potential scatter angles RHESSI can measure at any given instant. We modelled this effect by using the 0.15–2.0 MeV total count rate in the RHESSI instrument (Fig. 1) as the time-dependent flux template for a photon transport Monte Carlo simulation, and using the time-averaged GRB photon spectrum as measured by RHESSI for our input spectrum. This simulation used the detailed RHESSI mass model that has been developed under CERN's GEANT package, allowing us to model the instrument response to a GRB at the IPN⁸ sky coordinates for each rotation angle and instantaneous flux, assuming an unpolarized source. This distribution is also presented in the top panel of Fig. 2. We are looking for a modulation signal relative to this variation induced by the GRB time profile.

In the bottom panel of Fig. 2 we show the residual of the measured distribution once we have subtracted away the simulated response for an unpolarized GRB, showing our absolute modulation signal. For an unpolarized source we would expect this distribution to be flat, which we can rule out at an extremely high confidence level ($\chi^2 = 83.5$, 11 degrees of freedom, d.f.). When we fit this with a modulation curve, the fit improves significantly ($\chi^2 = 16.9$, 9 d.f.), with an amplitude of 128 ± 16 counts per bin. Statistically, this is a reasonable fit to the data (the probability of $\chi^2 > 16.9$ is 5%), but could be improved with even more detailed Monte Carlo simulations including time-dependent spectral variability. Using the count rates given above and the simulated distribution for an unpolarized GRB, we performed further numerical simulations to determine the probability that an unpolarized GRB could produce a modulation as large as the one we measure due to random Poisson counting statistics. We found this probability to be very low, $< 10^{-8}$, which translates to a confidence that we have measured a polarization at a level $> 5.7\sigma$. Finally, we estimated the modulation factor to be $\mu_m = 0.19 \pm 0.04$, using both a separate photon transport code which fully treats polarization in scattering and uses a simplified mass model, as well as analytical estimates based on the GEANT simulation with the full RHESSI mass model. Combining the modulation amplitude, the total source scatter event rate, and the RHESSI modulation factor, we derive a measured polarization $\Pi_m = 80 \pm 20\%$.

A number of tests were performed to check that the measured modulation is real. First, we verified that the simulated variation induced by the GRB light curve is accurate by comparing it to an angular distribution of events that were chance coincidences in two detectors. These interactions are nearly simultaneous, but separated by enough time to distinguish them as chance coincidences, not real scattered photons. This distribution should exhibit the same variations owing to the GRB light curve, but no polarization. When we subtracted the simulated distribution from the chance-coincident distribution, we found no evidence for a residual modulation. We have performed a number of independent checks to make sure we do not see modulations from other sources as well. We have verified that extended RHESSI background observations show no sign of modulations. In addition, we have done a preliminary analysis of a strong solar γ -ray flare observed on 23 July 2002, where we see some evidence for a modulation, but corresponding to a polarization $< 10\%$. Therefore we feel confident that we have characterized the systematic effects in RHESSI to below the 10% polarization level.

Received 7 February; accepted 12 March 2003; doi:10.1038/nature01612.

- Mészáros, P. Theories of gamma-ray bursts. *Annu. Rev. Astron. Astrophys.* **40**, 137–169 (2002).
- van Paradijs, J., Kouveliotou, C. & Wijers, R. A. M. J. Gamma-ray burst afterglows. *Annu. Rev. Astron. Astrophys.* **38**, 379–425 (2000).
- Lin, R. P. et al. The Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) mission. *Sol. Phys.* (in the press).
- Novick, R. Stellar and solar X-ray polarimetry. *Space Sci. Rev.* **18**, 389–408 (1975).
- Lei, F., Dean, A. J. & Hills, A. G. Compton polarimetry in gamma-ray astronomy. *Space Sci. Rev.* **82**, 309–388 (1997).
- McConnell, M. et al. in *Gamma Ray Bursts, 3rd Huntsville Symposium* (eds Kouveliotou, C., Briggs, M. F. & Fishman, G. J.) AIP Conf. Proc. **384**, 851–855 (1996).
- Hurley, K. et al. *GCN Circ.* 1727 (2002).
- Hurley, K. et al. *GCN Circ.* 1728 (2002).
- Rybicki, G. B. & Lightman, A. P. *Radiative Processes in Astrophysics* 180–181 (JWS, New York, 1979).
- Blandford, R. & Eichler, D. Particle acceleration at astrophysical shocks—a theory of cosmic-ray origin. *Phys. Rep.* **154**, 1–75 (1987).
- Frail, D. A., Waxman, E. & Kulkarni, S. R. A 450 day light curve of the radio afterglow of GRB 970508: fireball calorimetry. *Astrophys. J.* **537**, 191–204 (2000).
- Freedman, D. & Waxman, E. On the energy of gamma-ray bursts. *Astrophys. J.* **547**, 922–928 (2001).
- Derishev, E. V., Kocharovsky, V. V. & Kocharovsky, V. I. Physical parameters and emission mechanism in gamma-ray bursts. *Astron. Astrophys.* **372**, 1071–1077 (2001).
- Guetta, D., Spada, M. & Waxman, E. Efficiency and spectrum of internal gamma-ray burst shocks. *Astrophys. J.* **557**, 399–407 (2001).
- Covino, S. et al. GRB 990510: linearly polarized radiation from a fireball. *Astron. Astrophys.* **348**, L1–L4 (1999).
- Wijers, R. A. M. J. et al. Detection of polarization in the afterglow of GRB 990510 with the ESO Very Large Telescope. *Astrophys. J.* **523**, L33–L36 (1999).
- Rol, E. et al. GRB 990712: first indication of polarization variability in a gamma-ray burst afterglow. *Astrophys. J.* **544**, 707–711 (2000).
- Bersier, D. et al. The strongly polarized afterglow of GRB 020405. *Astrophys. J.* **583**, L63–L66 (2003).
- Waxman, E. γ -ray burst afterglow: confirming the cosmological fireball model. *Astrophys. J.* **489**, L33–L36 (1997).

- Galama, T. J. et al. The effect of magnetic fields on γ -ray bursts inferred from multi-wavelength observations of the burst of 23 January 1999. *Nature* **398**, 394–399 (1999).
- Medvedev, M. K. & Loeb, A. Generation of magnetic fields in the relativistic shock of gamma-ray burst sources. *Astrophys. J.* **526**, 697–706 (1999).
- Sari, R., Narayan, R. & Piran, T. Cooling timescales and temporal structure of gamma-ray bursts. *Astrophys. J.* **473**, 204–218 (1996).
- Gruzinov, A. & Waxman, E. Gamma-ray burst afterglow: polarization and analytic light curves. *Astrophys. J.* **511**, 852–861 (1999).
- Rees, M. J. & Mészáros, P. Unsteady outflow models for cosmological gamma-ray bursts. *Astrophys. J.* **430**, L93–L96 (1994).
- Spruit, H. C., Daigne, F. & Drenkhahn, G. Large scale magnetic fields and their dissipation in GRB fireballs. *Astron. Astrophys.* **369**, 694–705 (2001).
- Thompson, C. A model of gamma-ray bursts. *Mon. Not. R. Astron. Soc.* **270**, 480–498 (1994).
- Mészáros, P. & Rees, M. J. Poynting jets from black holes and cosmological gamma-ray bursts. *Astrophys. J.* **482**, L29–L32 (1997).
- Blandford, R. D. & Znajek, R. L. Electromagnetic extraction of energy from Kerr black holes. *Mon. Not. R. Astron. Soc.* **179**, 433–456 (1977).
- Usov, V. V. Millisecond pulsars with extremely strong magnetic fields as a cosmological source of gamma-ray bursts. *Nature* **357**, 472–474 (1992).
- Gruzinov, A. Gamma-ray burst phenomenology, shock dynamo, and the first magnetic fields. *Astrophys. J.* **563**, L15–L18 (2001).

Acknowledgements We thank D. Smith for help in learning RHESSI data analysis and providing simulation support, K. Hurley for IPN data and references, R. Lin, E. Quataert, J. Arons, C. Matzner and I. Fisk for discussions, and especially the RHESSI team for making all of their data immediately available to the public at (<http://rhesdatacenter.ssl.berkeley.edu>).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.E.B. (boggs@ssl.berkeley.edu).

Experimental entanglement purification of arbitrary unknown states

Jian-Wei Pan, Sara Gasparoni, Rupert Ursin, Gregor Weihs & Anton Zeilinger

Institut für Experimentalphysik, Universität Wien, Boltzmanngasse 5, 1090 Wien, Austria

Distribution of entangled states between distant locations is essential for quantum communication^{1–3} over large distances. But owing to unavoidable decoherence in the quantum communication channel, the quality of entangled states generally decreases exponentially with the channel length. Entanglement purification^{4,5}—a way to extract a subset of states of high entanglement and high purity from a large set of less entangled states—is thus needed to overcome decoherence. Besides its important application in quantum communication, entanglement purification also plays a crucial role in error correction for quantum computation, because it can significantly increase the quality of logic operations between different qubits⁶. Here we demonstrate entanglement purification for general mixed states of polarization-entangled photons using only linear optics⁷. Typically, one photon pair of fidelity 92% could be obtained from two pairs, each of fidelity 75%. In our experiments, decoherence is overcome to the extent that the technique would achieve tolerable error rates for quantum repeaters in long-distance quantum communication⁸. Our results also imply that the requirement of high-accuracy logic operations in fault-tolerant quantum computation can be considerably relaxed⁹.

The resource of quantum entanglement has many important applications in quantum information processing (QIP). In quantum communication, the generation of entanglement between distant

locations is essential for the long-distance realization of quantum cryptography¹, dense coding² and quantum teleportation³. Meanwhile, quantum information protocols involving entanglement are a very useful tool for fault-tolerant quantum computation⁹.

So far, significant experimental progress has been achieved in small-scale realizations of quantum communication^{10–14} (up to a few tens of kilometres) and quantum computation¹⁵ (up to a few qubits). However, serious problems occur in bringing QIP to technologically useful scales. One of the most important problems is the unavoidable decoherence due to the coupling between the quantum system and the environment. For instance, because of the noise in the quantum communication channel, the quality of entanglement between two particles is degraded more and more the further they propagate. Yet, the implementation of any of the above quantum communication schemes over large distances requires that two distant parties share highly entangled pairs. It is therefore necessary to overcome the unfavourable decoherence in any realistic large-scale realization of QIP.

One of the main tools used to overcome decoherence in QIP is entanglement purification^{4,5}—a method by which one can extract a smaller number of highly entangled pairs out of a large number of less-entangled pairs using only local operations and classical communication. Quantum repeaters⁸, based on both entanglement purification^{4,5} and entanglement swapping¹⁶, can thus provide an efficient solution to the problems of decoherence and photon loss in long-distance quantum communication¹⁰. Moreover, a recent study shows that entanglement purification is also important for fault-tolerant quantum computation because it can be used to increase the quality of logic operations between two qubits by several orders of magnitude⁶.

Recently, experimental efforts^{17–19} have been made to overcome some special decoherence processes using the ideas of local filtering

and entanglement concentration²⁰. However, the implementation of a general entanglement purification scheme^{4,5} that works for arbitrary unknown decoherence processes remains an experimental challenge. The main practical drawback of the original purification scheme^{4,5} is that it requires the CNOT operation. However, there is at present no implementation of CNOT gates between independent qubits that could be used for purification in the context of long-distance quantum communication. This is because the tolerable error rates of the CNOT operation must not exceed a few per cent⁸, which is far beyond what is experimentally possible at present. Fortunately, it was shown recently that this problem can be overcome by a general purification method that requires only linear optical elements⁷.

We now briefly explain the linear optical purification scheme⁷ (shown in Fig. 1) by discussing a specific example. Suppose that two distant parties, Alice and Bob, need to share photon pairs in the polarization entangled state $|\Phi^+\rangle_{ab}$ for a certain quantum communication task. We use the four usual Bell states:

$$|\Phi^\pm\rangle_{ab} = \frac{1}{\sqrt{2}}(|H\rangle_a|H\rangle_b \pm |V\rangle_a|V\rangle_b)$$

$$|\Psi^\pm\rangle_{ab} = \frac{1}{\sqrt{2}}(|H\rangle_a|V\rangle_b \pm |V\rangle_a|H\rangle_b)$$
(1)

where H (or V) denotes horizontal (or vertical) linear polarization, and subscripts a (or b) indicates the particle at Alice's (or Bob's) locations. Suppose further that, owing to noise in the quantum communication channel, the pairs they share before purification are in the mixed state:

$$\rho_{ab} = F|\Phi^+\rangle_{ab}\langle\Phi^+| + (1-F)|\Psi^-\rangle_{ab}\langle\Psi^-|$$
(2)

where $|\Psi^-\rangle_{ab}$ is an unwanted admixture. Without loss of generality, we choose this specific form of ρ_{ab} just for the convenience of discussion. The entanglement fidelity with respect to $|\Phi^+\rangle_{ab}$ is then given by $F = \langle\Phi^+|\rho_{ab}|\Phi^+\rangle$.

Alice and Bob start by choosing two pairs from the ensemble ρ_{ab} . Each of them then superimposes their photons on a polarizing beam splitter (PBS). The PBS with two input modes and two output modes transmits horizontal and reflects vertical polarization. An essential step in our purification scheme is to select those cases where there is exactly one photon in each of the four spatial output modes, which we refer to as 'four-mode cases'.

From equation (2) it follows that the original state of the two pairs can be seen as a probabilistic mixture of four cases: with a probability of F^2 , pairs 1 and 2 are in the state $|\Phi^+\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$, with equal probabilities of $F(1-F)$ in the states $|\Psi^-\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$ and $|\Phi^+\rangle_{a1b1} \cdot |\Psi^-\rangle_{a2b2}$, and with a probability of $(1-F)^2$ in $|\Psi^-\rangle_{a1b1} \cdot |\Psi^-\rangle_{a2b2}$.

It is easy to see that the cross combinations $|\Psi^-\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$ and $|\Phi^+\rangle_{a1b1} \cdot |\Psi^-\rangle_{a2b2}$ never lead to four-mode cases, because in these two cases only three photons always have equal polarization. Thus, by selecting only four-mode cases one can eliminate the cases that have one single bit-flip error.

Consider now the other two combinations $|\Phi^+\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$ and $|\Psi^-\rangle_{a1b1} \cdot |\Psi^-\rangle_{a2b2}$. Let us first discuss the $|\Phi^+\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$ case. In Fig. 1 we show that following our protocol Alice and Bob will get the state $|\Phi^+\rangle_{a3b3}$ whenever there is exactly one photon in each output mode, that is, with a probability of 50%. In the $|\Psi^-\rangle_{a1b1} \cdot |\Psi^-\rangle_{a2b2}$ case, following the same procedure, Alice and Bob will project the remaining two photons $a3$ and $b3$ into the state $|\Psi^+\rangle_{a3b3}$ with a probability of 50%.

Because the probabilities for $|\Phi^+\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$ and $|\Psi^-\rangle_{a1b1} \cdot |\Psi^-\rangle_{a2b2}$ are F^2 and $(1-F)^2$ respectively, after performing the purification procedure Alice and Bob will obtain the state $|\Phi^+\rangle_{a3b3}$ with a probability of $F^2/2$ and the state $|\Psi^+\rangle_{a3b3}$ with a probability of $(1-F)^2/2$. By applying our procedure, they can thus

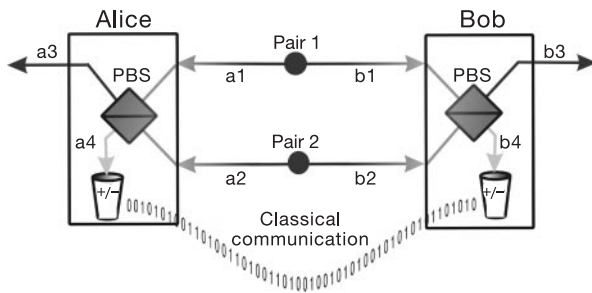


Figure 1 Schematic drawing showing the principle of entanglement purification using linear optics. We start with two less entangled pairs shared by Alice and Bob who superpose their photons on a polarizing beam splitter (PBS). Alice and Bob keep only those cases where there is exactly one photon in each output mode ('four-mode cases'). They perform a polarization measurement in the $\pm$ basis in modes $a4$ and $b4$, where $|+\rangle = (1/\sqrt{2})(|H\rangle + |V\rangle)$ and $|-\rangle = (1/\sqrt{2})(|H\rangle - |V\rangle)$. Depending on the results, Alice performs a specific operation on the photon in mode $a3$. After this procedure, the remaining pair in modes $a3$ and $b3$ will have a higher degree of entanglement than the two original pairs. To explain the protocol in detail, we consider the case where both pairs are in the state $|\Phi^+\rangle_{ab}$, which occurs with a probability of F^2 in our example. For the state $|\Phi^+\rangle_{a1b1} \cdot |\Phi^+\rangle_{a2b2}$, considering only those cases for which one, and only one, photon is finally found in modes $a4$ and $b4$ one obtains the state $(1/2)(|H\rangle_{a3}|H\rangle_{a4}|H\rangle_{b3}|H\rangle_{b4} + |V\rangle_{a3}|V\rangle_{a4}|V\rangle_{b3}|V\rangle_{b4})$. This shows that the probability for a four-mode case is 50%. Alice and Bob can then generate maximal two-photon entanglement between the output modes $a3$ and $b3$ out of the four-photon entanglement by performing polarization measurements on each of the two photons at $a4$ and $b4$ in the $\pm$ basis and comparing their results. If the measurement results at $a4$ and $b4$ are the same—that is, $|+\rangle|+\rangle$ or $|-\rangle|-\rangle$ —then the remaining two photons at $a3$ and $b3$ are left in the state $|\Phi^+\rangle_{a3b3}$. If the results are opposite, namely $|+\rangle|-\rangle$ or $|-\rangle|+\rangle$, then the remaining two photons are left in the state $|\Phi^-\rangle_{a3b3}$. In the second case, Alice can simply apply a local phase-flip operation on her remaining photon to convert the state $|\Phi^-\rangle_{a3b3}$ back to $|\Phi^+\rangle_{a3b3}$.

create a new ensemble described by the density operator

$$\rho'_{ab} = F'|\Phi^+\rangle_{ab}\langle\Phi^+| + (1 - F')|\Psi^-\rangle_{ab}\langle\Psi^-| \quad (3)$$

with a larger fraction of $F' = F^2/[F^2 + (1 - F)^2] > F$ (for $F > 1/2$) of pairs in the desired state $|\Phi^+\rangle_{ab}$ than before the purification.

Until now, it seems that we have only discussed a rather special example, that is, how to eliminate single bit-flip errors. However, it has been shown^{4,7} that with twirling (that is, rotating bases between individual purification steps or randomly choosing bases), the same method would also apply to the general mixed states ρ_{ab} , provided that they contain a sufficiently large fraction $F > 1/2$ of photon pairs in a maximally entangled state, $|\Phi^+\rangle_{ab}$ in our case. In short, this can be understood as follows. Using our purification method, one can first purify away single bit-flip errors. Phase errors can then be transformed into bit-flip errors by a 45° polarization rotation and treated in a subsequent purification step. Therefore, to demonstrate the generality of our scheme, it is sufficient to verify the purification

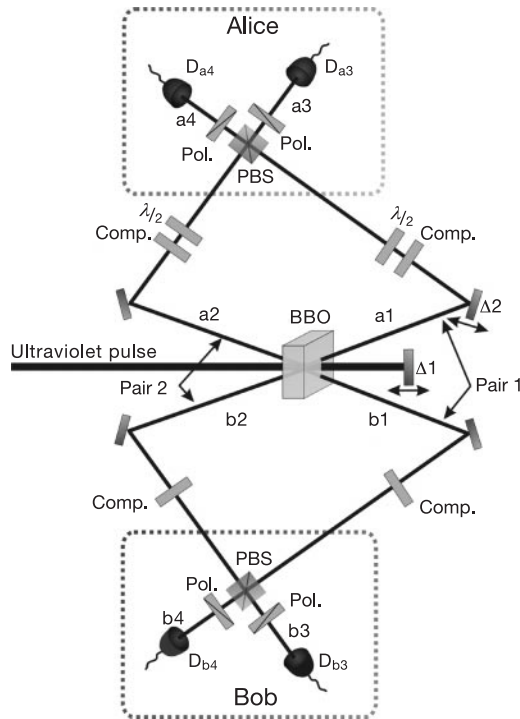


Figure 2 Experimental set-up for entanglement purification. A pulse of ultraviolet light passes through a BBO crystal twice to produce two polarization-entangled photon pairs, that is, pair 1 in a1–b1 and pair 2 in a2–b2. The ultraviolet laser with a central wavelength of 394 nm has a pulse duration of 200 fs and a repetition rate of 76 MHz. Four compensators (Comp.) are used to offset the birefringent effect caused by the BBO crystal during parametric down-conversion. In the experiment, with an average pump power of 500 mW, we are able to observe about 1.7×10^4 entangled pairs per second. The two photon pairs are originally prepared in the state $|\Phi^+\rangle_{ab}$, with a high signal-to-noise ratio of 30:1 in the $\pm$ basis. One member of each pair (the photons a1 and a2) is further sent through a half-wave plate ($\lambda/2$), whose angle is randomly set at either $+\delta$ or $-\delta$, to prepare the mixed state (2). Here, in order for the two pairs a1–b1 and a2–b2 to have roughly the same twofold coincidence, we have carefully chosen four detectors which have almost-equal detection efficiency. This arrangement ensures that the difference in twofold coincidence between pair a1–b1 and pair a2–b2 is less than 5% in all measurements. These features allow us to perform a precise analysis on the accuracy of our purification protocol. We then send the two pairs through the corresponding PBS to perform entanglement purification. By adjusting the positions of the delay mirrors $\Delta 1$ and $\Delta 2$, we can achieve simultaneous arrival of the photons at their respective PBS. Detecting exactly one photon in each of the four outputs (a3, a4, b3 and b4) behind a 45° polarizer (Pol.), one can verify the success of the purification scheme.

effect for the mixed state described in equation (2).

To demonstrate the purification scheme experimentally using only linear optical elements, we need to process two photon pairs which are in the mixed state (2). A schematic drawing of our experimental set-up is shown in Fig. 2. In our experiment, the required photon pairs are produced by parametric down-conversion from an ultraviolet pulsed laser in a beta barium borate (BBO) crystal²¹. We then interfere the two photons at Alice's (or Bob's) side in the pair a1–a2 (or b1–b2) at a PBS by making them indistinguishable. To achieve this indistinguishability, several methods have been used to guarantee that the photons at the same PBS have a perfect spatial and temporal overlap²² (see Fig. 3a and Methods section). After the four photons' passage through the two PBS, the

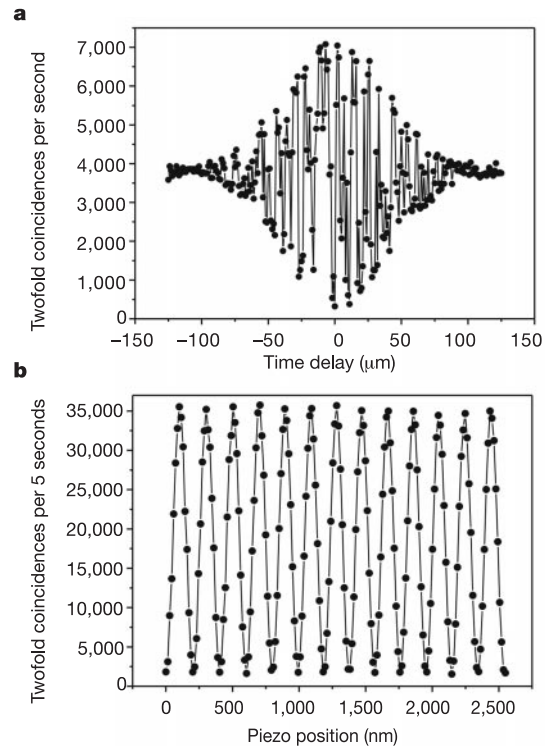


Figure 3 Experimental results showing the procedures to achieve perfect temporal overlap and to adjust the phase $\phi_4 = 0$. **a**, After roughly achieving the temporal overlap of modes b1 and b2, and of modes a1 and a2, we measure the twofold coincidence between the output modes a4 and b4 behind 45° polarizers, by scanning the position of $\Delta 1$ with a step size of $1 \mu\text{m}$. The envelope of the observed twofold coincidence varies indicating the visibility of the two-photon coherence. Outside the coherent region, $|H\rangle_{a4}|H\rangle_{b4}$ and $|V\rangle_{a4}|V\rangle_{b4}$ are distinguishable, so no interference occurs. Inside the coherent region, the best visibility is obtained at the position where perfect temporal overlap is achieved. We perform fine adjustment of the position of $\Delta 2$ and repeat the scanning of $\Delta 1$ until the best visibility is obtained. **b**, We use a piezo translation stage to move the mirror $\Delta 1$ to perform a fine scan around the region of zero delay, that is, around the centre of the envelope. By setting the piezo system to a position where we observe maximum twofold coincidence of a4–b4, ϕ_2 thus equals 0 and so does ϕ_4 . Note that in the experiment we also insert one 45° polarizer into each of the outputs a3 and b3 and measure their twofold coincidence. Another sine curve is observed (not shown). Usually the twofold coincidences of a4–b4 and of a3–b3 do not vary synchronously as we move the piezo system. This is caused by the birefringent effect, that is, the H and V polarizations of a photon accumulate a different phase during its passage through the PBS. This ultimately implies that the $|H\rangle|H\rangle$ and $|V\rangle|V\rangle$ components in a3–b3 acquire a relative phase ϕ'_2 , which is different from ϕ_2 . Experimentally, we introduce an additional birefringent element to compensate for the above effect. After this compensation, the two curves exhibit perfect synchronization, that is, $\phi_2 = \phi'_2$.

two photons in the mode pair a3–b3 have a higher probability to be in the desired state $|\Phi^+\rangle_{ab}$ if we detect one and only one photon polarized along the $\pm$ basis in each of the modes a4 and b4. Here we note that, owing to the absence of single-photon detectors, it is for the time being necessary also to detect the purified photons in a3 and b3 to ensure a fourfold event.

Thus far, we have considered only the ideal case, where there is at most one pair of photons in each of the input modes a1–b1 and a2–b2. However, with a probability of the same order of magnitude, two photon pairs will be emitted into the same mode pair because parametric down-conversion is a spontaneous process. This case could also result in four-mode cases in the real experiment. Specifically, if the two photon pairs are both in the mode pair a1–b1, we will then sometimes obtain a four-mode contribution that is in the state $|V\rangle_{a3}|H\rangle_{a4}|V\rangle_{b3}|H\rangle_{b4}$; and, if the two photon pairs are both in a2–b2, we will sometimes obtain the four-mode contribution $|H\rangle_{a3}|V\rangle_{a4}|H\rangle_{b3}|V\rangle_{b4}$.

As pointed out in ref. 7, if the position of the reflection mirror $\Delta 1$ is fixed and the amplitudes of these two four-mode contributions arrive at the two PBS simultaneously, then the two amplitudes will have a fixed relative phase (denoted by ϕ_4) and thus be in a coherent superposition $|H\rangle_{a3}|V\rangle_{a4}|V\rangle_{b3}|H\rangle_{b4} + e^{i\phi_4}|V\rangle_{a3}|H\rangle_{a4}|V\rangle_{b3}|H\rangle_{b4}$. Interestingly, adjusting the position of $\Delta 1$ such that $\phi_4 = 0$ and further performing a polarization measurement in the mode pair a4–b4 in the $\pm$ basis, in our purification protocol the two remaining photons in a3–b3 will always be converted into the state $|\Phi^+\rangle_{ab}$, which is exactly the desired maximally entangled state. This implies that the presence of double pair emission into the same mode pair in no way prevents us from carrying out the purification protocol, but rather makes the scheme more efficient (see ref. 23 for a detailed analysis).

Having achieved perfect spatial and temporal overlap and having fixed the relative phase ϕ_4 to zero (see Fig. 3 and the Methods section), we then experimentally demonstrated entanglement purification. In the first purification experiment, we prepare ρ_{ab} in the mixed state of equation (2) with an entanglement fidelity of $F = 0.75$, by randomly setting the half-wave plate axis to be oriented at $\pm 14^\circ$ (that is, $\delta = 14^\circ$). This implies that with a probability of 75%, the photon pairs will be in the desired state $|\Phi^+\rangle_{ab}$, and with a probability of 25% in the unwanted state $|\Psi^-\rangle_{ab}$. This is confirmed by the experimentally measured fractions both in the H/V and in the $\pm$ bases for the original mixed state, as shown in Fig. 4a and b respectively.

After purification, provided there is one and only one single photon detection in each of the modes a4 and b4 along the $\pm$ basis, we expect the two remaining photons in the modes a3 and b3 to acquire a high quality of entanglement. To verify this, we first measure the fractions in the H/V basis for the purified mixed state of a3–b3. The integration time is about 0.5 h for each of the four components HH , HV , VH and VV , and we collect about 700 fourfold coincidences for the maximum (HH or VV) and 60 for the minimum (HV or VH). The experimental results are shown in Fig. 4c. Compared with Fig. 4a, the fractions in Fig. 4c clearly confirm the significant improvement in the purity of the mixed state in the H/V basis.

Showing the improvement of purity in the H/V basis alone is a necessary but not a sufficient experimental criterion for the verification of entanglement purification, because the result in Fig. 4c is, in principle, both compatible with ρ'_{ab} (highly entangled) and with a statistical mixture of HH , HV , VH and HH (no entanglement at all). Thus, to exclude the latter case, we also measure the fractions of the purified state in the $\pm$ basis. Together with Fig. 4b, the results in Fig. 4d confirm the enhancement of entanglement well. Figure 4c and d show the significant improvement in entanglement fidelity, thus demonstrating the success of entanglement purification. The fidelity F' of the purified sub-ensemble is now about 0.92 ± 0.01 .

As a further demonstration, in a second experiment we also

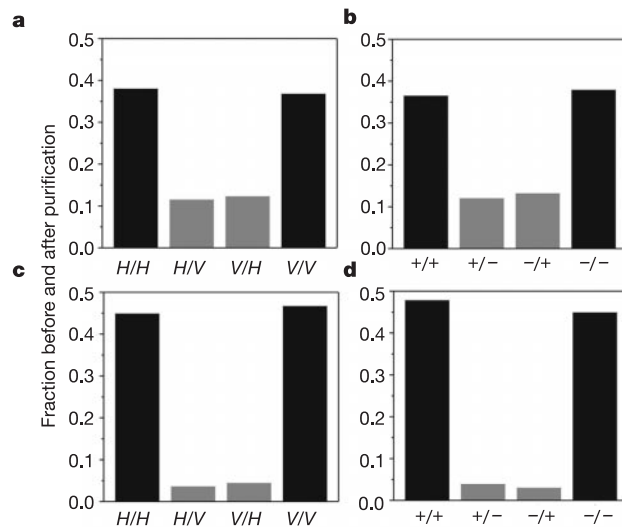


Figure 4 Experimental results. **a** and **b** show the experimentally measured fractions both in the H/V and in the $\pm$ bases for the original mixed state. **c** and **d** show the measured fractions of the purified state in the modes a3 and b3 both in the H/V and in the $\pm$ bases. Compared with the fractions in **a** and **b**, our experimental results shown in **c** and **d** together confirm the success of entanglement purification through the observed reduction of the fractions of the mixed (H/V and $\pm$) cases.

performed entanglement purification for the mixed state with fidelity parameter $F = 0.80$. After performing purification, the observed entanglement fidelity for the sub-ensemble in the modes a3 and b3 is about 0.94 ± 0.01 . This again verifies our purification scheme. It is worth noting that in both experiments, the visibilities of the original mixed states are 50% and 60%, respectively. However, after one-step purification the visibilities of the new ensembles are 84% and 88%: these are both well above the threshold above which Bell's inequality is violated, and are thus sufficient to ensure secure quantum communication.

We now briefly analyse the experimental accuracy of our purification method. As mentioned above, in our experiment we have two cases which contribute to the fourfold coincidence, that is, the case where there is one and only one entangled pair in each of the mode pairs a1–b1 and a2–b2, and the case where both entangled pairs are in the same mode pair. Consider an ensemble that is initially prepared in the mixed state (2). Thus, after purification the first case will result in a new mixed state that is described by equation (3), whereas the second case will always result in the desired entangled state $|\Phi^+\rangle_{ab}$. Note that these two cases contribute to the fourfold coincidence with the same probability. To determine the accuracy of our method, we first measure the entanglement fidelity of the pairs in a3–b3 (conditioned on a $|+\rangle|+\rangle$ coincidence detection in a4–b4) before introducing the channel noise. The observed fidelity is about 0.95, which represents the entanglement quality of the pairs in a3–b3 contributed from the double pair emission.

After taking into account this imperfection and subtracting the contribution of the double pair emission, we can calculate the purification accuracy of the first case. We consider, for example, the first experiment, in which the original ratio between the desired state and the unwanted state is 3:1. After purification, the final ratio of the first case is found to be about 8:1, which is in excellent agreement with the theoretical value 9:1 (refer to the text after equation (3)). From the obtained final ratio we can further estimate the accuracy of local operations at the PBS, which is better than 98%, or equivalently an error probability of at most 2%.

In the present experiment, we report the first demonstration of a general entanglement purification protocol. Entanglement purifi-

cation with high accuracy is important not only for quantum communication, but also for quantum computation. On the one hand, together with our recent experimental realization of high-fidelity teleportation¹⁴, the present experiment implies that the threshold of tolerable error rates in quantum repeaters can be achieved⁸. In that sense, we claim that our experimental results demonstrate for the first time the feasibility of overcoming the decoherence in scalable quantum communication. This opens up the possibility of realistic quantum communication over large distances. On the other hand, with the help of entanglement purification, the strict accuracy requirements of the gate operations for fault-tolerant quantum computation can also be significantly reduced⁶.

The methods developed in our purification experiment have many useful applications in the field of linear optics QIP and in experimental tests of quantum nonlocality. It was noted recently that, while our set-up directly enables an implementation of a non-destructive C-NOT gate with a success probability of 25% (ref. 24), a slight modification of the set-up also provides a simple way to implement a nonlinear sign gate²⁵—a fundamental element in linear optics quantum computation²⁶. Furthermore, our methods of achieving perfect spatial-temporal overlap and phase stabilization provide the necessary techniques for experimental investigations of schemes in quantum communication with linear optics and atomic ensembles²⁷.

Meanwhile, although the requirement of phase stabilization appears to be a drawback in long-distance quantum communication, this problem can be solved in our purification scheme by using entangled photon pairs produced with quantum dots²⁸ instead of by parametric down-conversion. Finally, the two-photon four-dimensional entanglement exploited in the experiment also enables us to perform an experimental test of ‘all versus nothing’ quantum nonlocality for two particles²⁹ and high-efficiency entanglement-assisted quantum cryptography³⁰. □

Methods

Making two independent photons indistinguishable

To make two independent photons indistinguishable, one has to guarantee that the two photons have good spatial and temporal overlap at the PBS. To achieve this, the two outputs of the PBS are spectrally filtered (3-nm bandwidth) and monitored by fibre-coupled detectors (either D_{a3} and D_{a4} or D_{b3} and D_{b4}). While the single-mode fibre couplers act as spatial filters to guarantee good mode overlap of the detected photons, the narrow bandwidth filters stretch the coherence time to about 700 fs, substantially larger than the pump-pulse duration. The filtering process effectively erases any possibility of distinguishing the two photons according to their arrival time and therefore leads to interference²².

To meet the condition of temporal overlap, the two photons at the same location (Alice’s or Bob’s) must arrive at their PBS simultaneously. To achieve this, we first move $\Delta 1$ in small steps to search for the position (denoted by p_b) where the two photons in b1 and b2 have the same arrival time. Then we insert one 45° polarizer into each of the modes b1 and b2 and observe the two-photon Hong–Ou–Mandel (HOM) dip after the PBS by performing polarization measurements in both modes b3 and b4 in the $\pm$ basis. The maximum interference occurs at the region of zero delay, that is, at the centre of the HOM dip. In a similar manner, we can find the position of $\Delta 1$ (denoted by p_a) where the two photons in a1 and a2 have the same arrival time. Fixing the position of $\Delta 1$ at p_b and moving the mirror $\Delta 2$ by a distance $p_b - p_a$, we can thus roughly achieve temporal overlap. After this preliminary alignment, we take away the polarizers in the input modes. Because the polarizers have slightly different time delays, it is necessary to further improve the temporal overlap both of the mode pair a1–a2 and of the mode pair b1–b2.

Here we exploit a two-photon four-dimensional entangled state^{23,29}. This corresponds to the case where one and only one entangled pair is created after the pump pulse has passed through the BBO crystal twice, that is, the entangled photon pair is emitted into a superposition of the mode pairs a1–b1 and a2–b2. Thus, before the photon pair passes through the noisy channel the two-photon four-dimensional entangled state can be written as $(|H\rangle_a|H\rangle_b + |V\rangle_a|V\rangle_b)(|a1\rangle|b1\rangle + e^{i\phi_2}|a2\rangle|b2\rangle)$, where $|a1\rangle|b1\rangle$ and $|a2\rangle|b2\rangle$ denote the spatial wavefunctions of the photon pair, and their relative phase is ϕ_2 determined by the position of the reflection mirror $\Delta 1$. Furthermore, if the amplitudes of the spatial modes a1–a2 and b1–b2 arrive at their own PBS simultaneously, then after the four-dimensional entangled state passes through the PBS, the two photons in a4–b4 will be in the polarization state $|H\rangle_{a4}|H\rangle_{b4} + e^{i\phi_2}|V\rangle_{a4}|V\rangle_{b4}$. By observing two-photon interference fringes between the modes a4 and b4, we can then achieve perfect temporal overlap (see Fig. 3a for details).

Achieving a fixed phase of $\phi_4 = 0$

Usually, one can find the right position of $\Delta 1$, that is, $\phi_4 = 0$, by measuring the fourfold polarization correlation in the $\pm$ basis. However, in the real experiment, owing to the low fourfold coincidence rate, which was about five per second on average, it is very difficult to find the correct position of $\Delta 1$ in this way. To achieve $\phi_4 = 0$, here we again exploit the two-photon four-dimensional entangled state. Because the two relative phases ϕ_4 and ϕ_2 satisfy the relation $\phi_4 = 2\phi_2$, it follows that if $\phi_2 = 0$, then $\phi_4 = 0$ (see ref. 23). Indeed, we can then determine the relative phase by measuring, for example, the polarization correlation for the two photons in the mode pair a4–b4. Therefore, by inserting a 45° polarizer in each of the output modes a4 and b4 and scanning the reflection mirror $\Delta 1$ we can observe a two-photon interference sine curve in the mode pair a4–b4. Setting $\Delta 1$ to a position where a maximum coincidence rate for $|+\rangle|+\rangle$ is obtained, we can fix both phases ϕ_4 and ϕ_2 to 0.

To have a fixed phase of $\phi_4 = 0$, phase stabilization is required throughout the whole measurement. To achieve this, we built the set-up on a thick aluminium platform to avoid sound resonance, used feedback air-conditioning to avoid thermal expansion of the interferometer, and built a plastic housing around the set-up to avoid air flow. At the same time, we also use the twofold coincidence of a4–b4 as a phase monitor. In this way, we are able to keep the phase stable for several hours. We emphasize that, as opposed to the earlier multi-photon experiments, here we require interferometric precision and stability where the scale is the wavelength and is therefore much smaller than coherence stability. This was particularly difficult and challenging to achieve.

Received 17 January; accepted 31 March 2003; doi:10.1038/nature01623.

- Ekert, A. Quantum cryptography based on Bell’s theorem. *Phys. Rev. Lett.* **67**, 661–663 (1991).
- Bennett, C. H. & Wiesner, S. J. Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states. *Phys. Rev. Lett.* **69**, 2881–2884 (1992).
- Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **83**, 3081–3084 (1993).
- Bennett, C. H. *et al.* Purification of noisy entanglement, and faithful teleportation via noisy channels. *Phys. Rev. Lett.* **76**, 722–725 (1996).
- Deutsch, D. *et al.* Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Phys. Rev. Lett.* **77**, 2818–2821 (1996).
- Duer, W. & Briegel, H.-J. Entanglement purification for quantum computation. *Phys. Rev. Lett.* **90**, 067901 (2003).
- Pan, J.-W., Simon, C., Brukner, C. & Zeilinger, A. Entanglement purification for quantum communication. *Nature* **410**, 1067–1070 (2001).
- Briegel, H.-J., Duer, W., Cirac, J. I. & Zoller, P. Quantum repeaters: The role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998).
- Gottesman, D. & Chuang, I. L. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature* **402**, 390–393 (1999).
- Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. Quantum cryptography. *Rev. Mod. Phys.* **74**, 145–195 (2002).
- Mattle, K., Weinfurter, H., Kwiat, P. G. & Zeilinger, A. Dense coding in experimental quantum communication. *Phys. Rev. Lett.* **76**, 4656–4659 (1996).
- Bouwmeester, D. *et al.* Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
- Marcikic, I., de Riedmatten, H., Tittel, W., Zbinden, H. & Gisin, N. Long-distance teleportation of qubits at telecommunication wavelengths. *Nature* **421**, 509–513 (2003).
- Pan, J.-W., Gasparoni, S., Aspelmeyer, M., Jennewein, T. & Zeilinger, A. Experimental realization of freely propagating teleported qubits. *Nature* **421**, 721–725 (2003).
- Jones, J. Quantum computing: Putting it into practice. *Nature* **421**, 28–29 (2003).
- Zukowski, M., Zeilinger, A., Horne, M. A. & Ekert, A. “Event-ready-detectors” Bell experiment via entanglement swapping. *Phys. Rev. Lett.* **71**, 4287–4290 (1993).
- Kwiat, P. G., Barzaza-Lopez, S., Stefanov, A. & Gisin, N. Experimental entanglement distillation and ‘hidden’ non-locality. *Nature* **409**, 1014–1017 (2001).
- Yamamoto, T., Koashi, M., Ozdemir, S. K. & Imoto, N. Experimental extraction of an entangled photon pair from two identically decohered pairs. *Nature* **421**, 343–346 (2003).
- Zhao, Z., Yang, T., Chen, Y.-A., Zhang, A.-N. & Pan, J.-W. Experimental realization of entanglement concentration and a quantum repeater. *Phys. Rev. Lett.* (in the press); preprint available at (<http://xxx.lanl.gov/quant-ph/0211075>) (2003).
- Bennett, C. H., Bernstein, H. J., Popescu, S. & Schumacher, B. Concentrating partial entanglement by local operations. *Phys. Rev. A* **53**, 2046–2052 (1996).
- Kwiat, P. G. *et al.* New high intensity source of polarization-entangled photon. *Phys. Rev. Lett.* **75**, 4337–4341 (1995).
- Zukowski, M., Zeilinger, A. & Weinfurter, H. Entangling photons radiated by independent pulsed source. *Ann. NY Acad. Sci.* **755**, 91–102 (1995).
- Simon, C. & Pan, J.-W. Polarization entanglement purification using spatial entanglement. *Phys. Rev. Lett.* **89**, 257901 (2002).
- Pittman, T. B., Jacobs, B. C. & Franson, J. D. Demonstration of nondeterministic quantum logic operations using linear optical elements. *Phys. Rev. Lett.* **88**, 257902 (2002).
- Rudolph, T. & Pan, J.-W. A simple gate for linear optics quantum computing. Preprint available at (<http://xxx.lanl.gov/quant-ph/0108056>) (2001).
- Knill, E., Laflamme, R. & Milburn, G. J. A scheme for efficient quantum computation with linear optics. *Nature* **409**, 46–52 (2001).
- Duan, L.-M., Lukin, M. D., Cirac, J. I. & Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414**, 413–418 (2001).
- Santori, C., Fattal, D., Vuckovic, J., Solomon, G. S. & Yamamoto, Y. Indistinguishable photons from a single-photon device. *Nature* **419**, 594–596 (2001).
- Chen, Z.-B., Pan, J.-W., Zhang, Y.-D., Brukner, C. & Zeilinger, A. All-versus-nothing violation of local realism for two entangled photons. *Phys. Rev. Lett.* **90**, 160408 (2003).
- Zhao, Z., Yang, T., Chen, Z.-B., Du, J.-F. & Pan, J.-W. Deterministic and highly efficient quantum cryptography with entangled photon pairs. Preprint available at (<http://xxx.lanl.gov/quant-ph/0211098>) (2002).

Acknowledgements We thank H. Briegel, T. Jennewein and C. Simon for discussions. This work was supported by the Austrian Science Foundation (FWF), and by the TMR and the QuComm programs of the European Commission.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.-W. P. (pan@ap.univie.ac.at) or A.Z. (zeilinger-office@exp.univie.ac.at).

Real-time detection of electron tunnelling in a quantum dot

Wei Lu^{*†}, Zhongqing Ji^{*}, Loren Pfeiffer[‡], K. W. West[‡] & A. J. Rimberg^{*§}

^{*} Department of Physics and Astronomy, § Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77005, USA

[‡] Bell Laboratories, Lucent Technologies, Inc., Murray Hill, New Jersey 07974, USA

Nanostructures in which strong (Coulomb) interactions exist between electrons are predicted to exhibit temporal electronic correlations¹. Although there is ample experimental evidence that such correlations exist², electron dynamics in engineered nanostructures have been observed directly only on long time-scales³. The faster dynamics associated with electrical currents or charge fluctuations⁴ are usually inferred from direct (or quasi-direct) current measurements. Recently, interest in electron dynamics has risen, in part owing to the realization that additional information about electronic interactions can be found in the shot noise⁵ or higher statistical moments^{6,7} of a direct current. Furthermore, interest in quantum computation has stimulated investigation of quantum bit (qubit) readout techniques^{8,9}, which for many condensed-matter systems ultimately reduce to single-shot measurements of individual electronic charges. Here we report real-time observation of individual electron tunnelling events in a quantum dot using an integrated radio-frequency single-electron transistor^{10,11}. We use electron counting to measure directly the quantum dot's tunnelling rate and the occupational probabilities of its charge state. Our results provide evidence in favour of long (10 μ s or more) inelastic scattering times in nearly isolated dots.

Real-time detection of individual electrons is a formidable task: the electronic charge is small, and typical timescales for electronic dynamics are short, ranging from picoseconds to microseconds. We therefore require a detector with both a low charge noise of $\delta q \approx 1 \times 10^{-5} e \text{ Hz}^{-1/2}$ and a fast response time of roughly a microsecond or less. The recently developed radio-frequency single-electron transistor (RF-SET)^{10,11} satisfies both requirements, and we use such a device as the basis of our detection scheme. Our system of choice for investigation of charge dynamics is a quantum dot (QD), a small conducting region in a semiconductor that contains a few to a few thousand electrons. Sufficiently small QDs contain well-defined energy levels, and are often referred to as artificial atoms. A QD is typically connected to macroscopic leads by tunnel barriers, allowing access to its internal state. Importantly, these tunnel barriers and other dot properties can be adjusted to allow studies of electron dynamics.

We have performed measurements on three different samples, each consisting of a QD and an integrated RF-SET detector. All three

showed similar behaviour, and here we present results from two of them (S1 and S2). When the QD is formed by application of gate voltages (Fig. 1a), it contains a well-defined number of electrons N ; at low temperatures and bias voltages N can change only by ± 1 . The capacitive SET–QD coupling ensures that such a change shifts the polarization charge Q_{SET} of the SET island by some fraction of an electronic charge e ; typically $\Delta Q_{\text{SET}} \approx (0.1\text{--}0.2)e$ for our samples. The shift ΔQ_{SET} changes the differential resistance R_d of the SET, which in turn allows SET-based electrometry^{12,13}. In radio-frequency (r.f.) operation, changes in R_d modulate the amplitude of a carrier wave reflected from a resonant circuit containing the SET (Fig. 1b). Demodulating the carrier wave recovers the modulating signal (Fig. 1c), forming the basis for real-time measurements of the QD charge Q_{QD} .

When the dot is sufficiently isolated from the surrounding two-dimensional electron gas (2DEG), we observe switching behaviour in the RF-SET output (Fig. 1d) reminiscent of random telegraph signals (RTSs) associated with engineered³ or naturally occurring^{14,15} charge traps. We have taken pains to ensure that the RTS we observe is due to individual electron tunnelling events on the QD. The size of the RTS for S1 as determined by comparison to a 0.05e r.m.s. sine wave (Fig. 1c and d) corresponds to $\Delta Q_{\text{SET}} \approx 0.1e$,

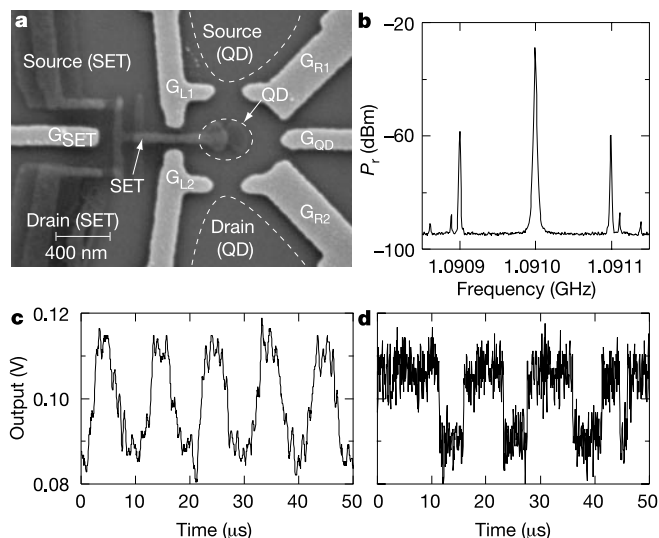


Figure 1 Characterization of RF-SET response. **a**, Electron micrograph of the sample design for S1. We begin with a GaAs/AlGaAs heterostructure containing a 2DEG located 190 nm below the sample surface. At 4 K, the 2DEG sheet density is $1.3 \times 10^{11} \text{ cm}^{-2}$ and its mobility is $4.1 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. The SET consists of a small Al island connected to a source and drain through small tunnel barriers (total resistance 15 k Ω) formed by a thin AlO_x layer and has a charging energy $E_{C_{\text{SET}}} = e^2/2C_{\text{SET}} \approx 162 \mu\text{eV}$ where C_{SET} is the total SET capacitance. The SET is superconducting for all measurements, which were made in a dilution refrigerator at its base temperature of 15 mK. The dot is formed by applying a negative voltage to the Au gates G_{L1} , G_{L2} , G_{R1} , G_{R2} and G_{QD} , leaving a small pool of electrons at their centre. The SET island extends between G_{L1} and G_{L2} to lie above the QD and maximize sensitivity to its charge. Gate G_{SET} is used to adjust the SET offset charge Q_{SET} . The QD can be coupled by tunnel barriers to its own source and drain, or isolated completely from them. The dot has an estimated area $A \approx 200 \text{ nm} \times 300 \text{ nm}$ and contains roughly 80 electrons with an average level spacing $\Delta \approx 1/Ag_{2D} \approx 60 \mu\text{eV}$; here $g_{2D} = m^*/\pi \hbar^2$ is the two-dimensional density of states, and m^* is the effective mass of GaAs. **b**, Power spectral density P_r of a reflected carrier wave at 1.091 GHz for which Q_{SET} is modulated by a 100-kHz, 0.05e r.m.s. sine wave. For S1, a d.c. SET bias of 500 μV and r.f. amplitude of 87 μV r.m.s. gave a charge sensitivity of $\delta q \approx 2.4 \times 10^{-5} e \text{ Hz}^{-1/2}$ that was relatively insensitive to Q_{SET} . **c**, Demodulated signal for the reflected wave in **b**, passed through a low-pass filter (12 dB per octave, 1 MHz corner frequency) and sampled by a digital oscilloscope. **d**, RTS observed when the dot gates are sufficiently energized that the QD tunnelling rate Γ lies within our 1-MHz bandwidth.

[†] Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

the same as that from d.c. measurements. The RTS appears only when all the dot gates are energized to form a QD; charge traps would cause switching without dot formation. Finally, as discussed below, the RTS varies periodically with an applied gate voltage with the same period as that determined by standard d.c. electrometry (Fig. 2).

We use a time-domain analysis of the RTS to count the number of

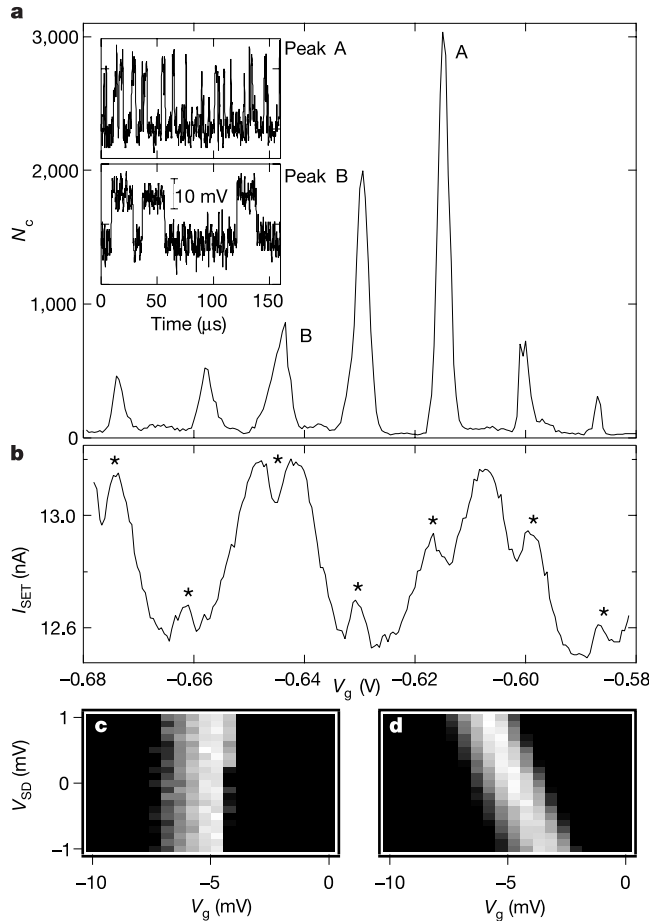


Figure 2 Time-domain analysis of the RF-SET output for S1. **a**, Number of counts N_c detected by the RF-SET versus voltage V_g applied to G_{DD} at zero dot bias. For each value of V_g , a span $\Delta t = 16$ ms is recorded with a digital oscilloscope, sampled at 40 ns per point for a total of 400,000 points. The bin size $M \approx 200$ is chosen to minimize noise (favouring large M) while allowing at most one transition per bin (favouring small M); N_c is then nearly independent of the threshold η . N_c varies periodically with V_g , behaviour extremely unlikely to arise from charge traps near the SET. A clear variation in tunnelling rates from peak to peak is visible (insets). At the peaks the RF-SET output spends roughly equal time in each state. **b**, Measured d.c. SET current versus V_g . Varying V_g causes both the SET and dot charges Q_{SET} and Q_{DD} to change. Q_{DD} varies quickly owing to its stronger coupling C_g to G_{DD} and causes the small, rapid oscillations (marked by asterisks). Direct variations in Q_{SET} cause the slow oscillations. The QD period is 15 mV ($C_g \approx 11$ aF), the same as for N_c , while the SET period is 38 mV. From the change in SET current due to $\Delta N = \pm 1$, we estimate $\Delta Q_{SET} \approx 0.1e$. **c**, Greyscale image of N_c versus V_g and a symmetrically applied source-drain bias V_{SD} (that is, $V_S = +V_{SD}/2$ and $V_D = -V_{SD}/2$) near a degeneracy point. Independence of N_c from V_{SD} indicates complete isolation from the leads. As the QD offset charge can be expressed as $Q_{DD} = C_g V_g + C_S V_S + C_D V_D$, it also indicates that $C_S \approx C_D$, where C_S and C_D are the capacitances of the QD to its source and drain, respectively. **d**, Greyscale image of N_c versus V_g and an asymmetric bias (that is, $V_S = V_{SD}$ and $V_D = 0$). The location of the degeneracy point (but not N_c) now varies with V_{SD} . Using $C_g \approx 11$ aF, we obtain $C_S \approx C_D \approx 20$ aF. If we estimate that the total QD capacitance is about 200 aF, then the QD charging energy is $E_{C_{DD}} \approx 0.4$ meV. For $\Gamma = 2.32 \times 10^{-5} \text{ s}^{-1}$, the dimensionless conductance $g \approx 1.6 \times 10^{-5}$.

transitions N_c (either upward or downward) in the RF-SET output. Within a bin of M sampled points we test the likelihood ratio of two hypotheses: that an abrupt change between two mean values did or did not occur at some point k . For white gaussian noise, the likelihood ratio test is¹⁶:

$$\max_k [k\mu_0^2 - (M-k)\mu_1^2 - M\mu^2] > \eta \quad (1)$$

where $\mu_0 = k^{-1} \sum_{i=1}^k x_i$, $\mu_1 = (M-k)^{-1} \sum_{i=k+1}^M x_i$, $\mu = M^{-1} \sum_{i=1}^M x_i$, the x_i are the data values and η is a threshold. If the test succeeds, a change is said to have taken place at the point k for which the maximum occurred. The bin is then shifted over the entire data range in a given time trace. We find that N_c varies periodically with applied gate voltage V_g (Fig. 2a) and also that its maximum value differs from peak to peak. The latter variations arise from changes in the QD tunnelling rate $\Gamma = \Delta g/h$ (insets), where Δ is the level spacing and g the dimensionless conductance (in units of e^2/h) determined by the quantum mechanical coupling between a given dot level and the leads. We verify that the RTS is due to the QD charge by using the SET as a d.c. electrometer (Fig. 2b). The SET current is multiply periodic in V_g , characteristic of coupled SET-QD systems¹⁷. The faster oscillations, corresponding to changes in Q_{DD} , have the same periodicity as N_c . The counting data therefore represent a direct measurement of equilibrium charge fluctuations on a QD; the peaks in N_c correspond to charge degeneracy points at which the dot charge changes from N to $N+1$.

For sample S1, we were unable to perform transport measurements on the QD while in the strongly isolated regime in which we detected switching events, as we could not completely close the channel between G_{L1} and G_{L2} (Fig. 1a) while maintaining stable SET operation. Instead, we increased the voltage applied to G_{R1} and G_{R2} so that the dot was isolated from its source and drain, and coupled to the 2DEG only through the channel to its left. Isolation from source and drain was verified both by d.c. measurements and the more sensitive charge counting possible with the RF-SET. For the latter, both N_c and the location of the charge degeneracy point are independent of a symmetric bias V_{SD} (Fig. 2c). An asymmetric bias (Fig. 2d) only shifts the location of the degeneracy point due to changes in Q_{DD} , again leaving N_c unaffected. The dot is therefore electrically isolated from its source and drain, and symmetrically placed between them. As the QD is connected to the 2DEG by a single tunnel barrier, analysis near a degeneracy point is simplified.

Nearly isolated dots such as that in S1 have been of great interest lately^{18,19}, owing to predictions that inelastic scattering processes within them will be strongly affected by the level spacing Δ (ref. 20). The QDs in both S1 and S2 are in the quantum regime for which individual dot levels are important: $k_B T < \Delta < E_{C_{DD}}$, where k_B is Boltzmann's constant. We use a master equation approach²¹ to analyse the charge dynamics for S1, beginning with a model in which electrons tunnel to and from a single quantum level on the dot. Labelling the states with and without an extra electron $|1\rangle$ and $|0\rangle$ respectively, detailed balance gives:

$$\Gamma P_0 f(\Delta U) = \Gamma P_1 [1 - f(\Delta U)] \quad (2)$$

where f is the Fermi function, P_0 and P_1 are the occupational probabilities of $|0\rangle$ and $|1\rangle$, $\Delta U = 2E_{C_{DD}} \left(\frac{1}{2} - \frac{Q_{DD}}{e} \right) + E_1$, $E_{C_{DD}}$ is the QD charging energy and E_1 is the dot level referred to the Fermi energy. P_0 and P_1 are given by:

$$P_0 = 1 - f(\Delta U) \text{ and } P_1 = f(\Delta U) \quad (3)$$

and the tunnelling rates between the states by $\Gamma_{0 \rightarrow 1} = \Gamma f(\Delta U)$ and $\Gamma_{1 \rightarrow 0} = \Gamma [1 - f(\Delta U)]$.

Comparing this model with our data (Fig. 3a), we find that P_0 and P_1 are well described by equation (3) with an effective 2DEG electron temperature $T_{\text{eff}} \approx 0.43$ K. This elevated temperature, comparable to results for charge traps¹⁵, is unsurprising, because the SET is immediately above the 2DEG. Comparing $\Gamma_{0 \rightarrow 1}$ and

$\Gamma_{1 \rightarrow 0}$ with experimental values (Fig. 3b), we again see excellent agreement. No additional fitting parameters are required, as the value of Γ used ($\sim 2.32 \times 10^5 \text{ s}^{-1}$) is determined experimentally. Electron counting has allowed us to measure directly the occupational probabilities of the QD charge state as a function of Q_{QD} , as well as its tunnelling rate Γ to the 2DEG.

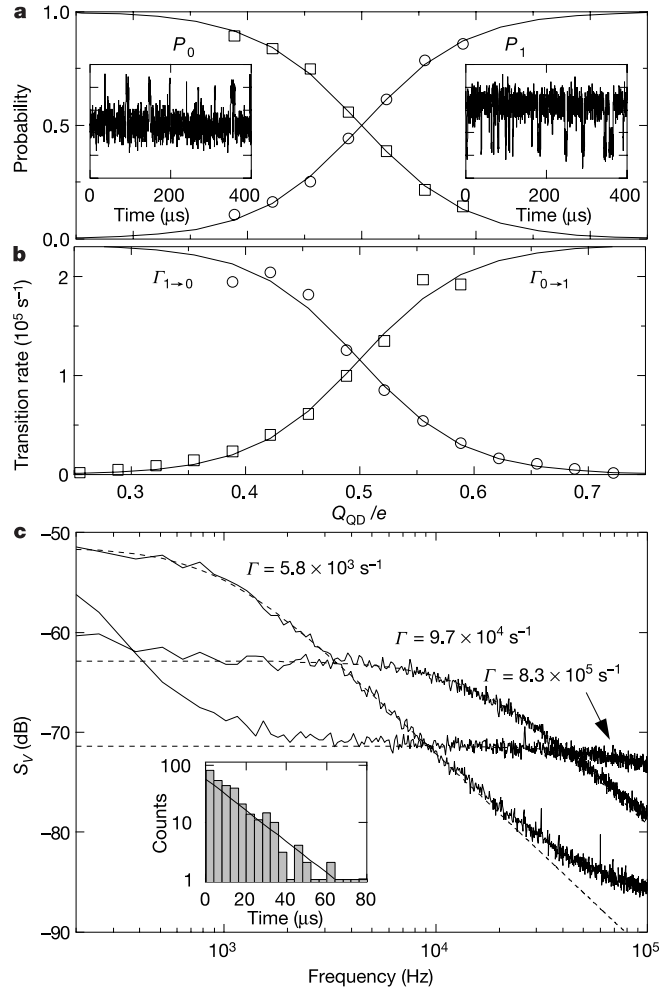


Figure 3 Comparison of a simple one-state model with the RTS near a charge degeneracy point for S1. **a**, Calculated probability of occupation of $|0\rangle$ and $|1\rangle$. Once transitions are located in a time trace, they are used to calculate the average times τ_0 and τ_1 spent in $|0\rangle$ and $|1\rangle$, respectively. Then $P_0 = \tau_0/(\tau_0 + \tau_1)$ and $P_1 = \tau_1/(\tau_0 + \tau_1)$, allowing a direct measurement of P_0 and P_1 . We measured a series of 15 time traces for different Q_{QD} ; the RF-SET output shifted from being primarily high to primarily low (insets) on opposite sides of the degeneracy point. For the central seven data points, we could accurately measure both τ_0 and τ_1 and therefore P_0 (squares) and P_1 (circles). Farther from degeneracy the lifetime of the energetically unfavourable state becomes too short to be accurately measured. Treating E_1 as the zero of energy, we perform a fit (lines) using equation (3) to the central seven points from which we extract an effective temperature T_{eff} based on our estimated E_{CQD} . **b**, For a poissonian process $\Gamma_{0 \rightarrow 1} = \tau_0^{-1}$ (squares) and $\Gamma_{1 \rightarrow 0} = \tau_1^{-1}$ (circles), and we can calculate the total tunnelling rate from $\Gamma = \tau_0^{-1} + \tau_1^{-1}$. The average of $\tau_0^{-1} + \tau_1^{-1}$ for the points in **a** gives a measured value of Γ used to calculate $\Gamma_{0 \rightarrow 1}$ and $\Gamma_{1 \rightarrow 0}$ (lines) as described in the text. Agreement is excellent, even for points for which only one of τ_0 and τ_1 could be measured. **c**, Power spectra of the RF-SET output near three different degeneracy points with different tunnelling rates. Dashed lines are fits to a lorentzian function, from which we extract the tunnelling rates shown in the figure. We can also fit the distribution of times in either $|0\rangle$ or $|1\rangle$ to an exponential (solid line), as shown for $|0\rangle$ in the inset for the spectrum with $\Gamma = 9.7 \times 10^4 \text{ s}^{-1}$. The time $15.8 \mu\text{s}$ from this fit is in excellent agreement with the mean $\tau_0 = 15.6 \mu\text{s}$ and close to the standard deviation of $14.3 \mu\text{s}$, another indication of a poissonian process.

If the tunnelling is a poissonian process, for which consecutive events are uncorrelated, we expect the power spectral density $S_V(f)$ at the RF-SET output voltage to have a lorentzian form $S_V(f) \propto 1/[\Gamma^2 + (2\pi f)^2]$, (ref. 14), in excellent agreement with measured spectra (Fig. 3c). Values of Γ from a fit to this form agreed exactly with results of time-domain analysis. Finally, we find that the time the QD spends in either $|0\rangle$ or $|1\rangle$ is exponentially distributed (inset, Fig. 3c), another indication of a poissonian process. The equilibrium charge fluctuations on the QD are thus well described by tunnelling to and from a single quantum level, governed by the single timescale $1/\Gamma$.

The significance of these results relates to inelastic transitions in isolated dots, which (when arising from electron–electron interactions) have been predicted to vanish below a temperature T_c parametrically greater²⁰ than Δ ; in our case $T_c \approx 1.9\Delta/k_B \approx 1.3 \text{ K}$. A master equation was later used to address inelastic transitions arising from other sources, for example, external radiation¹⁹. Within this framework, a total inelastic scattering rate Γ_{in} (comparable to Γ) would lead to tunnelling between multiple dot levels and the 2DEG. The probabilities P_0 and P_1 for each dot level i obey detailed balance (equation (2)). However, the total probabilities $\sum_i P_0$ and $\sum_i P_1$ of the dot containing zero or one extra electron do not obey a simple relation, unless the tunnelling rates Γ_i for all levels are identical. Thus, either $\Gamma_{\text{in}} \ll \Gamma$, or multi-level tunnelling dynamics can be described by a single timescale. Although simplicity and

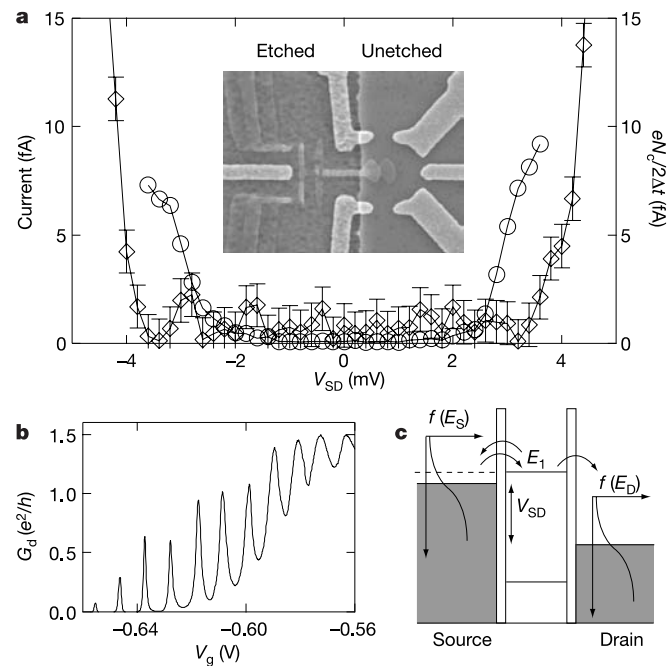


Figure 4 Measurement of non-equilibrium QD charge fluctuations for S2. The total SET resistance was $58 \text{ k}\Omega$, its charging energy $237 \mu\text{eV}$ and its charge sensitivity $\delta q \approx 2.2 \times 10^{-5} e \text{ Hz}^{-1/2}$. **a**, Comparison of d.c. current I measured using a current amplifier with $eN_c/2\Delta t$. The dot was held at a fixed offset charge Q_{QD} and a symmetrically applied source drain voltage V_{SD} was increased in $200 \mu\text{V}$ steps. For each value of V_{SD} , 40 values of I were measured and averaged, and their absolute value plotted (diamonds). A time trace was also taken and N_c calculated (circles). The modified sample (inset) allowed these measurements to be performed as switching could be observed for significantly smaller voltages on G_{R1} and G_{R2} . **b**, Coulomb-blockade oscillations in the QD differential conductance G_d , indicating that the etch does not prevent dot formation. This measurement was performed with significantly stronger source–drain coupling than used for the counting measurements. **c**, Energy level diagram when the source Fermi level is just below an unoccupied dot level. An electron that enters the dot from the source may either leave through the drain or return to the source. The relative likelihood of the two possibilities depends on the values of Γ_S and Γ_D .

agreement with recent theory²⁰ make the former possibility attractive, we cannot rule out the latter without comparison to a more complete calculation of charge noise in QDs.

We have also examined non-equilibrium charge fluctuations in a second sample (S2). We modified our design by removing the 2DEG immediately to the left of the dot with a wet etch (inset, Fig. 4a). Doing so does not prevent formation of a QD (Fig. 4b), and allows us to observe switching due to tunnelling between the dot and its source and drain. Both a time-domain analysis and d.c. current measurement (Fig. 4a) show a wide Coulomb-blockade region near zero bias. As the bias increases, the counts N_c detected by the RF-SET rise first, followed later by a detectable current I . This is in agreement with the single level model discussed above, which gives $I \approx e\Gamma_S\Gamma_D f(E_S)/(\Gamma_S + \Gamma_D)$ when, for example, the source is close to resonance with an unoccupied dot level (Fig. 4c). Here Γ_S and Γ_D are the couplings between the dot and its source and drain and $E_S = 2E_{CQD}(\frac{1}{2} - \frac{Q_{QD}}{e}) + \frac{1}{2}eV_{SD}$. The corresponding expression for N_c scaled to units of current is $N_c e/2\Delta t = e\Gamma_S f(E_S)\{\Gamma_D + [1 - f(E_S)]\Gamma_S\}/(\Gamma_S + \Gamma_D)$; the factor of two on the left accounts for the two transitions an electron makes entering and leaving the dot. The current senses only electrons that traverse the dot. In contrast, N_c includes electrons that enter and leave the dot through the same junction (second term in curly brackets), that is, the RF-SET also detects thermal charge fluctuations at the source junction (Fig. 4c). If $\Gamma_S \approx \Gamma_D$, the two terms make roughly equal contributions to N_c just below threshold, and $N_c e/2\Delta t > I$ as measured.

Received 16 January; accepted 7 April 2003; doi:10.1038/nature01642.

- Ben-Jacob, E. & Gefen, Y. New quantum oscillations in current driven small junctions. *Phys. Lett. A* **108**, 289–292 (1985).
- Delsing, P., Likharev, K. K., Kuzmin, L. S. & Claeson, T. Time-correlated single-electron tunneling in one-dimensional arrays of ultrasmall tunnel junctions. *Phys. Rev. Lett.* **63**, 1861–1864 (1989).
- Dresselhaus, P. D., Ji, L., Han, S., Lukens, J. E. & Likharev, K. K. Measurement of single electron lifetimes in a multijunction trap. *Phys. Rev. Lett.* **72**, 3226–3229 (1994).
- Berman, D., Zhitenev, N. B., Ashoori, R. C. & Shayegam, M. Observation of quantum fluctuations of charge on a quantum dot. *Phys. Rev. Lett.* **82**, 161–164 (1999).
- Blanter, Y. M. & Büttiker, M. Shot noise in mesoscopic conductors. *Phys. Rep.* **336**, 1–166 (2000).
- Levitov, L. S. & Lesovik, G. B. Charge-transport statistics in quantum conductors. *JETP Lett.* **55**, 555–559 (1992).
- Levitov, L. S., Lee, H. & Lesovik, G. B. Electron counting statistics and coherent states of electric current. *J. Math. Phys.* **37**, 4845–4866 (1996).
- Shnirman, A. & Schön, G. Quantum measurements performed with a single-electron transistor. *Phys. Rev. B* **57**, 15400–15407 (1998).
- Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57**, 120–126 (1998).
- Schoelkopf, R. J., Wahlgren, P., Kozhevnikov, A. A., Delsing, P. & Prober, D. E. The radio-frequency single-electron transistor (RF-SET): A fast and ultrasensitive electrometer. *Science* **280**, 1238–1242 (1998).
- Aassime, A., Johansson, G., Wendin, G., Schoelkopf, R. J. & Delsing, P. Radio-frequency single-electron transistor as readout device for qubits: Charge sensitivity and backaction. *Phys. Rev. Lett.* **86**, 3376–3379 (2001).
- Likharev, K. K. Single-electron transistors: Electrostatic analogs of the DC SQUIDS. *IEEE Trans. Magn.* **23**, 1142–1145 (1987).
- Fulton, T. A. & Dolan, G. J. Observation of single-electron charging effects in small tunnel junctions. *Phys. Rev. Lett.* **59**, 109–112 (1987).
- Kirton, M. J. & Uren, M. J. Noise in solid-state microstructures: A new perspective on individual defects, interface states, and low-frequency (1/f) noise. *Adv. Phys.* **38**, 367–468 (1989).
- Fujisawa, T. & Hirayama, Y. Charge noise analysis of an AlGaAs/GaAs quantum dot using transmission-type radio-frequency single-electron transistor technique. *Appl. Phys. Lett.* **77**, 543–545 (2000).
- Basseville, M. & Nikiforov, I. V. *Detection of Abrupt Changes* (Prentice Hall, Englewood Cliffs, 1993).
- Lu, W., Rimberg, A. J., Maranowski, K. D. & Gossard, A. C. A single-electron transistor strongly coupled to an electrostatically defined quantum dot. *Appl. Phys. Lett.* **77**, 2746–2748 (2000).
- Folk, J. A., Marcus, C. M. & Harris, J. S. Decoherence in nearly isolated quantum dots. *Phys. Rev. Lett.* **87**, 206802 (2001).
- Eisenberg, E., Held, K. & Altshuler, B. L. Dephasing times in closed quantum dots. *Phys. Rev. Lett.* **88**, 136801 (2002).
- Altshuler, B. L., Gefen, Y., Kamenev, A. & Levitov, L. S. Quasiparticle lifetime in a finite system: A nonperturbative approach. *Phys. Rev. Lett.* **78**, 2803–2806 (1997).
- Beenakker, C. W. J. Theory of Coulomb-blockade oscillations in the conductance of a quantum dot. *Phys. Rev. B* **44**, 1646–1656 (1991).

Acknowledgements We thank W. L. Wilson, M. Thalakulam, J. Sarkar, R. J. Schoelkopf, D. H. Johnson, D. Natelson, R. M. Westervelt, D. Driscoll and A. C. Gossard for discussions and experimental assistance. This work was supported by the National Science Foundation, the Army Research Office, and the Robert A. Welch Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.R. (rimberg@rice.edu).

Spin entropy as the likely source of enhanced thermopower in $\text{Na}_x\text{Co}_2\text{O}_4$

Yayu Wang*, Nyrissa S. Rogado†, R. J. Cava‡ & N. P. Ong*‡

* Department of Physics, † Department of Chemistry, ‡ Princeton Materials Institute, Princeton University, Princeton, New Jersey 08544, USA

In an electric field, the flow of electrons in a solid produces an entropy current in addition to the familiar charge current. This is the Peltier effect, and it underlies all thermoelectric refrigerators. The increased interest in thermoelectric cooling applications has led to a search for more efficient Peltier materials and to renewed theoretical investigation into how electron–electron interaction may enhance the thermopower of materials such as the transition-metal oxides^{1–4}. An important factor in this enhancement is the electronic spin entropy, which is predicted^{4–6} to dominate the entropy current. However, the crucial evidence for the spin-entropy term, namely its complete suppression in a longitudinal magnetic field, has not been reported until now. Here we report evidence for such suppression in the layered oxide $\text{Na}_x\text{Co}_2\text{O}_4$, from thermopower and magnetization measurements in both longitudinal and transverse magnetic fields. The strong dependence of thermopower on magnetic field provides a rare, unambiguous example of how strong electron–electron interaction effects can qualitatively alter electronic behaviour in a solid. We discuss the implications of our finding—that spin-entropy dominates the enhancement of thermopower in transition-metal oxides—for the search for better Peltier materials.

In a thermopower experiment, a temperature gradient $-\nabla T$ drives the diffusion of charge carriers to the cooler end of the sample. The charge accumulation leads to a net electric field E which determines the thermopower (or Seebeck coefficient) $Q = E/|\nabla T|$ (ref. 7). In conventional metals, the flow of electrons induced by $-\nabla T$ is nearly cancelled by a parallel flow of holes, so that Q is strongly reduced by the factor T/T_F from the ‘natural value’ $k_B/e \approx 86 \mu\text{V K}^{-1}$, where k_B is Boltzmann’s constant, e the electron charge and T_F the Fermi temperature⁷. Moreover, as the currents are indifferent to the spins, a longitudinal magnetic field $\mathbf{H} \parallel -\nabla T$ has almost no effect on Q . In materials with strong electron–electron interaction, however, the spin degrees are predicted to produce a large contribution of the Heikes form^{2,4,6}:

$$Q = \mu/eT = -\sigma/e \quad (1)$$

where μ is the chemical potential, and σ (the entropy per electron)

equals $k_B \ln(g_s g_c)$ with g_s and g_c the spin and configuration degeneracies, respectively. From equation (1), the enhancement to Q from spin entropy is of the order k_B/e , but it is suppressed to zero if the spin degeneracy is lifted in a magnetic field ($g_s \rightarrow 1$).

Q in $\text{Na}_x\text{Co}_2\text{O}_4$ (at 300 K) is about ten times larger than in typical metals⁸. The role of spin fluctuations, as in heavy-fermion systems, has been suggested⁸. Despite Hall effect⁸, heat capacity⁹ and nuclear spin-resonance¹⁰ experiments, however, the enhancement in Q remains puzzling. Experiments have not ruled out either strong-interaction theories involving spin configurations¹¹, or conventional interpretations based on small Fermi surfaces¹². By investigating the thermopower and magnetization in both longitudinal and transverse H , we show that $\text{Na}_x\text{Co}_2\text{O}_4$ is in fact a strong-correlation system in which the spin entropy term accounts for almost all of Q at 2 K and a dominant fraction at 300 K.

In $\text{Na}_x\text{Co}_2\text{O}_4$, the Co ions occupy the sites of a two-dimensional (2D) triangular lattice. Chemical arguments and band-structure calculations¹² show that the O $2p$ states lie far below the Co $3d$ states and the chemical potential falls within the band formed from t_{2g} states in Co. Hence the electrons donated by Na ions are distributed among the Co ions, a fraction (δ) of which are in the Co^{4+} state, while the rest ($1 - \delta$) are Co^{3+} . From $x = 1.36$ (determined by inductively coupled plasma spectrometer (ICP) chemical analysis of

our crystals), we find $\delta = 0.32$. The variation of the susceptibilities χ_{ab} and χ_c versus temperature T (upper inset in Fig. 1) fitted accurately to the Curie–Weiss form $1/(T + \theta)$, implying the existence of large local moments that are antiferromagnetically coupled. The plot of $1/\chi_{ab}$ gives $\theta = 55 \pm 5$ K and a moment around $0.87 \mu_B$ averaged over all Co ions (μ_B is the Bohr magneton). This experimentally constrains the Co^{4+} ions to be in the low-spin state, with spin $S = 1/2$ and moment $\sqrt{3} \mu_B$, while the Co^{3+} ions are in their low-spin $S = 0$ state (Fig. 1, lower inset), in agreement with ref. 10.

The in-plane resistivity ρ displays a T -linear variation below 100 K that is suggestive of strong-correlation behaviour (dashed line in Fig. 1). As noted in ref. 8, Q (circles in Fig. 1) increases with T to values much larger than in conventional metals. However, it is the effect of a magnetic field on Q that is the most informative. The pronounced effect of an in-plane field H on Q is shown in Fig. 2 ($H \parallel -\nabla T$). Stringent precautions against spurious contributions to Q were adopted (see Supplementary Information). Even at moderately high T (30 K), Q decreases significantly in a 14-T field. Lowering the temperature makes the field-suppression of Q more pronounced, until at 4 K, Q changes sign and saturates near 12 T to a ‘floor’ value equal to $-0.25 \mu\text{V K}^{-1}$. With further decrease in T , the floor value itself monotonically decreases to zero. The trace at 2.5 K shows that Q decreases monotonically to zero in a field of 8 T, and remains at zero over an extended field range.

The strong effect of field on Q is also observed with $H \parallel c$ (normal to plane), except that, above 6 K, we observe an additional term that produces an increase in Q in weak fields. At 4.2 K, Q displays the same field profile as above, but with a proportionately weaker decrease (Fig. 3). As the sample is warmed to 10 K and higher, the relative change in Q is initially positive in low fields. At higher fields,

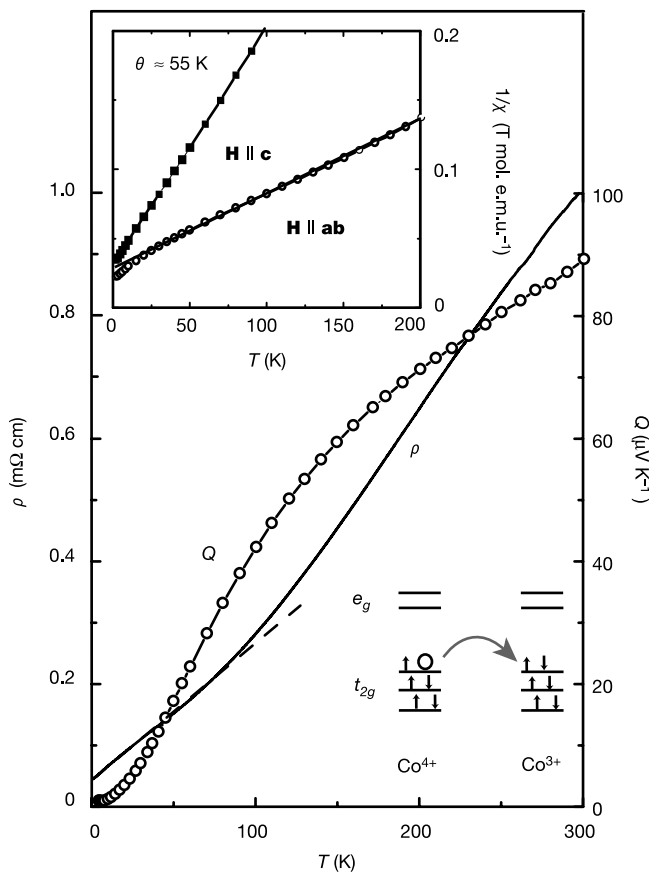


Figure 1 The temperature (T) dependence of magnetic and transport properties of single-crystal $\text{Na}_x\text{Co}_2\text{O}_4$ and electronic states in the Co ions. The in-plane thermopower Q (circles) and resistivity ρ (solid curve) are plotted versus T . The T -linear dependence of ρ below 100 K is indicated by the dashed line. The upper inset shows the T dependence of the inverse susceptibilities χ_{ab} and χ_c measured in a 5-T field applied in-plane and along the c axis, respectively. The lower inset is a sketch of the low-spin t_{2g} states in the Co^{4+} and Co^{3+} ions. The hopping of a hole (arrow) is accompanied by the transfer of a spin-1/2 excitation.

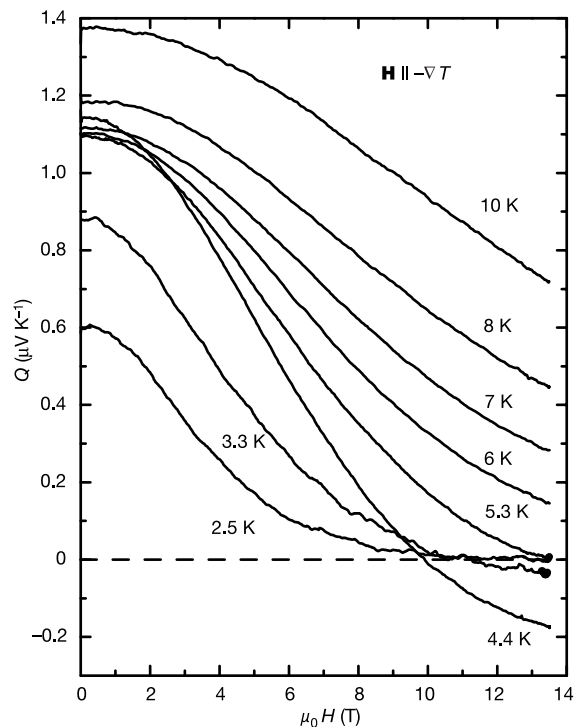


Figure 2 The in-plane thermopower Q versus an in-plane $H \parallel (-\nabla T)$ at selected T . At 4.4 K, Q tends asymptotically at high fields to a negative floor value ($-0.25 \mu\text{V K}^{-1}$). At the lowest T (2.5 K), Q approaches a value unresolved from zero when H exceeds 8 T. The applied temperature difference δT is less than 0.3 K below 4 K, and about 0.5 K above 4 K. By using phosphor bronze as the voltage leads, we have eliminated background contributions to Q (Supplementary Information).

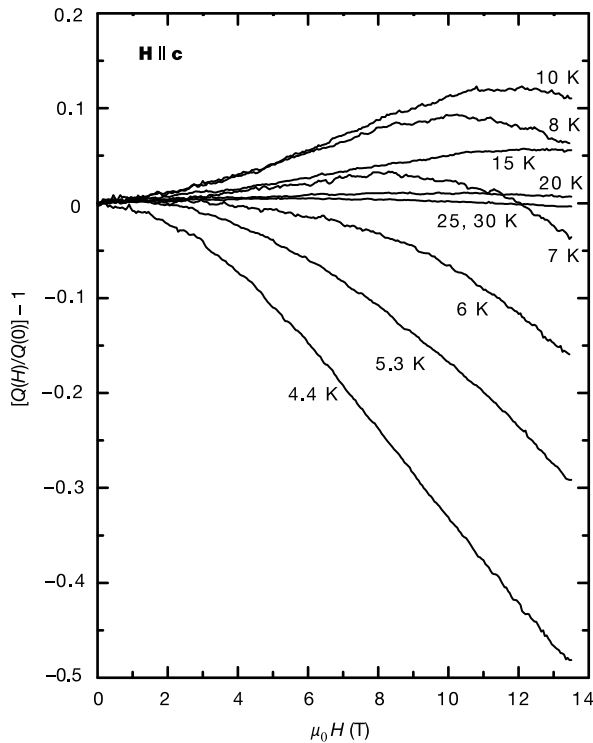


Figure 3 The relative change in Q versus a transverse field $H \parallel c$ in $\text{Na}_x\text{Co}_2\text{O}_4$. At 6 K and lower, Q decreases monotonically with H , but at higher temperatures, Q initially increases in weak field before decreasing at larger H .

the negative trend again becomes apparent. The overall pattern in transverse field is consistent with a positive contribution superposed on the term intrinsic to the spin degrees uncovered in the longitudinal field.

The evidence that the H dependence of Q arises from the spin degrees of freedom derives from both the transport and susceptibility experiments. Because an in-plane field couples to the spin degrees only, but not to the orbital degrees, the strong H dependence of Q in Fig. 3 must arise from the effect of field on the spins. In the geometry with $H \parallel c$, we reason that the non-monotonic field dependence of Q also arises entirely from the effect of H on the spin degrees, as the carrier mean-free path $l \approx 100 \text{ \AA}$ estimated from ρ is far too short to produce any orbital effect in a field of 14 T. Comparing curves at 4 K in Figs 2 and 3, we infer that a similar decrease in Q requires a c -axis field that is 1.4 times larger than the in-plane field. The close agreement between this ratio and the ratio of the bulk susceptibilities χ_{ab}/χ_c at 4 K strongly suggests that the field anisotropies in the magnetization and thermopower experiments have the same origin. The field suppression of Q is larger with the field in-plane compared with the field along the c axis because it is easier to align the spins with an in-plane field.

In conventional metals, a longitudinal field merely changes the relative populations of the spin-up (+) and spin-down (−) electrons without significantly altering the entropy current in each spin population⁷. The complete suppression displayed in Fig. 2 is incompatible with a conventional band picture.

We next discuss how the field suppression may be understood within a strong-correlation scenario. As shown in the lower inset of Fig. 4, in the 2D triangular lattice of the Co ions, a fraction $\delta = 1/3$ are in the Co^{4+} state in which $S = 1/2$, while the rest are in the Co^{3+} state with $S = 0$ (the higher-lying e_g orbitals are unoccupied). The

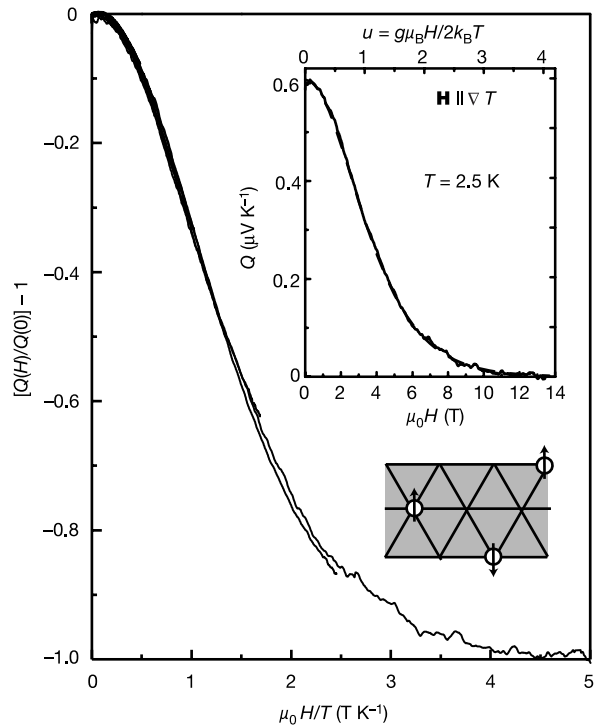


Figure 4 The suppression of the thermopower by an in-plane magnetic field in $\text{Na}_x\text{Co}_2\text{O}_4$. The normalized Q is plotted versus $\mu_0 H/T$. Curves measured at various T values (2.5, 6, 8, 10, 12.5, 15, 20 and 25 K) collapse to a universal curve. The lower inset shows the spin-1/2 excitations in an inert background of Co^{3+} ions. The upper inset compares the data curve at 2.5 K (solid curve) with the expression in equation (2) (dashed curve).

elementary charge-transport process is the hopping of a hole from Co^{4+} to Co^{3+} (Fig. 1, lower inset). A large on-site repulsion excludes double occupancy of a site by the holes. Because this process converts the Co^{4+} ($S = 1/2$) to a Co^{3+} ($S = 0$) and vice versa, we also transfer a spin-1/2 along with the hole, which implies a transfer of spin entropy $\sigma = k_B \ln 2$ in zero field. The excitations are holes carrying charge $+1e$ and spin-1/2 moving in a sea of Co^{3+} sites, which are magnetically and electrically inert. An E field leads to the heat current $J_Q = nvk_B T \ln 2$ as well as the charge current $J = nev$ (with n the hole concentration and v the average velocity). As the ratio J_Q/J (called the Peltier coefficient Π)⁷ equals the product QT , we obtain the spin-entropy part of Q given in equation (1). Moreover, because the spin-entropy current is in the direction of the charge current, this contribution to Q is positive (hole-like) as observed.

As T decreases below $\theta \approx 50 \text{ K}$ (the antiferromagnetic exchange energy $J_{AF} \approx \theta$), the spin entropy must decrease rapidly with incipient magnetic ordering to vanish at $T = 0$. When $T \ll \theta$, the majority of the spins become antiferromagnetically correlated with their neighbours, leaving only a small fraction T/θ to behave as free spins. In an E field, this residual population of free spins carries the spin entropy, although all the holes remain itinerant. This implies that the ratio J_Q/J decreases steeply below θ , consistent with the observed behaviour of Q in Fig. 1.

In a magnetic field, the residual spin entropy $\sigma(H, T)$ is further decreased, eventually reaching zero when H is strong enough completely to lift the twofold degeneracy. Modelling the residual free spins as non-interacting spins with the Landé g -factor g , we calculate the field dependence of the entropy as:

$$\sigma(H, T)/\sigma(0, T) = \{\ln[2 \cosh(u)] - u \tanh(u)\}/\ln(2) \quad (2)$$

where $u = g\mu_B H/2k_B T$. This dependence should apply to the component of Q derived from the spin entropy.

As shown in Fig. 4, the curves of Q in Fig. 2 between 2.5 and 25 K can be collapsed onto a universal curve when plotted against the ratio H/T . Using the curve at 2.5 K (which extends to the largest u), we find that equation (2) provides a remarkably close fit if the Landé g -factor is 2.2 ± 0.1 (upper inset in Fig. 4). We view the close fit as confirming evidence that the entropy current observed in the thermopower indeed derives entirely from the spin-1/2 excitations, and the complete suppression of Q reflects the removal of the spin degeneracy by the applied field.

We note that, at each T , the observed Q is composed of several distinct contributions. Application of a strong longitudinal H completely suppresses only the spin-entropy term, leaving the remaining terms unaffected (the configuration entropy $k_B \ln g_c$, for instance) which are revealed as the floor value at high fields. In Fig. 2, these terms amount to $-0.25 \mu\text{V K}^{-1}$ at 4.4 K and rapidly vanish at lower T . We emphasize that the spin entropy contribution to Q is not restricted to low T . Measurements at high fields (30 T) show that Q is suppressed by about 20% at 30 K. Assuming that, when $T \gg \theta$, the spin contribution asymptotes to $(k_B/e) \ln(2) \approx 60 \mu\text{V K}^{-1}$, we find that it constitutes about two-thirds of the total thermopower at 300 K, thus accounting for most of the enhanced thermopower.

The spin-entropy enhancement may be widespread in the transition-metal oxides. A key ingredient inferred from our measurements is the motion of charge tied to local moments of spin-1/2 in a background of Co^{3+} ions that are magnetically inert (diamagnetic). Of the large class of cobalt oxides, the Co ions are nearly always in their low-spin state. In the Co oxides that can be made conducting, therefore, the thermopower is likely to be dominated by spin-entropy terms, suggesting that this is a promising class of materials to search for improved Peltier materials.

Received 28 January; accepted 28 March 2003; doi:10.1038/nature01639.

- Mahan, G., Sales, B. & Sharp, J. Thermoelectric materials: New approaches to an old problem. *Phys. Today* **50**(3), 42–47 (1997).
- Beni, G. Thermoelectric power of the narrow-band Hubbard chain at arbitrary electron density: Atomic limit. *Phys. Rev. B* **10**, 2186–2189 (1974).
- Palsson, G. & Kotliar, G. Thermoelectric power near the density-driven Mott transition. *Phys. Rev. Lett.* **80**, 4775–4778 (1998).
- Chaikin, P. M. & Beni, G. Thermopower in the correlated hopping regime. *Phys. Rev. B* **13**, 647–651 (1976).
- Kwak, J. F. & Beni, G. Thermoelectric power of a Hubbard chain with arbitrary electron density: Strong-coupling limit. *Phys. Rev. B* **13**, 652–657 (1976).
- Kwak, J. F., Beni, G. & Chaikin, P. M. Thermoelectric power in Hubbard-model systems with different densities: *N*-methylphenazinium-tetracyanoquinodimethane (NMP-TCNQ), and quinolinium ditetracyanoquinodimethane. *Phys. Rev. B* **13**, 641–646 (1976).
- Ziman, J. M. *Principles of The Theory of Solids* 235 (Cambridge Univ. Press, London, 1972).
- Terasaki, I., Sasago, Y. & Uchinokura, K. Large thermoelectric power in NaCo_2O_4 single crystals. *Phys. Rev. B* **56**, R12685–R12687 (1997).
- Ando, Y., Miyamoto, N., Segawa, K., Kawata, T. & Terasaki, I. Specific-heat evidence for strong electron correlations in the thermoelectric material $(\text{Na,Ca})\text{Co}_2\text{O}_4$. *Phys. Rev. B* **60**, 10580–10583 (1999).
- Ray, R., Ghoshray, A., Ghoshray, K. & Nakamura, S. ^{59}Co NMR studies of metallic NaCo_2O_4 . *Phys. Rev. B* **59**, 9454–9461 (1999).
- Koshibae, W., Tsutsui, K. & Maekawa, S. Thermopower in cobalt oxides. *Phys. Rev. B* **62**, 6869–6872 (2000).
- Singh, D. J. Electronic structure of NaCo_2O_4 . *Phys. Rev. B* **61**, 13397–13402 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank S. Hannahs for technical assistance. We acknowledge support from the US National Science Foundation (NSF). Some of the measurements were performed at the US National High Magnetic Field Laboratory, which is supported by the NSF and the state of Florida.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N.P.O. (npo@princeton.edu).

¹⁴⁶Sm–¹⁴²Nd evidence from Isua metamorphosed sediments for early differentiation of the Earth's mantle

Guillaume Caro*, Bernard Bourdon*, Jean-Louis Bircck* & Stephen Moorbath†

* Laboratoire de Géochimie et Cosmochimie (UMR 7579 CNRS), Institut de Physique du Globe de Paris, Université Denis Diderot, 4 place Jussieu, 75252 Paris Cedex 05, France

† Department of Earth Sciences, Oxford University, Parks Road, Oxford OX1 3PR, UK

Application of the ¹⁴⁷Sm–¹⁴³Nd chronometer (half-life of 106 Gyr) suggests that large-scale differentiation of the Earth's mantle may have occurred during the first few hundred million years of its history¹. However, the signature of mantle depletion found in early Archaean rocks is often obscured by uncertainties resulting from open-system behaviour of the rocks during later high-grade metamorphic events². Hence, although strong hints exist regarding the presence of differentiated silicate reservoirs before 4.0 Gyr ago, both the nature and age of early mantle differentiation processes remain largely speculative^{3–5}. Here we apply short-lived ¹⁴⁶Sm–¹⁴²Nd chronometry (half-life of 103 Myr) to early Archaean rocks using ultraprecise measurement of Nd isotope ratios. The analysed samples are well-preserved metamorphosed sedimentary rocks from the 3.7–3.8-Gyr Isua greenstone belt of West Greenland. All display well-resolved ¹⁴²Nd anomalies averaging 15 ± 4 p.p.m. (2σ). Using the initial $\epsilon^{143}\text{Nd}$ value from ref. 6 coupled ¹⁴⁶Sm–¹⁴⁷Sm chronometry constrains the mean age of mantle differentiation to $4,460 \pm 115$ Myr. This early Sm/Nd fractionation probably reflects differentiation of the Earth's mantle during the final stage of terrestrial accretion.

Short-lived radioactive nuclides are ideal tools to constrain early differentiation of the main terrestrial reservoirs. The ¹²⁹I–¹²⁹Xe (half-life 16 Myr) chronometer constrains mantle degassing and formation of the atmosphere to the first 100 Myr of Earth history⁷, and ¹⁸²Hf–¹⁸²W systematics (half-life 9 Myr) provide clear evidence that core segregation occurred during the first 30 Myr of terrestrial accretion^{8–10}. Similarly, the short-lived ¹⁴⁶Sm–¹⁴²Nd chronometer has long been viewed as a promising tracer for exploring the early history of the Earth's mantle. The presence of live ¹⁴⁶Sm in the early Solar System was first demonstrated by Lugmair *et al.*¹¹, and its initial abundance subsequently defined as ¹⁴⁶Sm/¹⁴⁴Sm = 0.008 ± 1 at 4.566 Gyr ago¹². Because of its low initial abundance, ¹⁴⁶Sm is effectively extinct after 4–5 half-lives, so that ¹⁴²Nd/¹⁴⁴Nd anomalies can be related solely to differentiation of silicate reservoirs during the first few hundred million years of Earth history. Additionally, the ¹⁴²Nd signature of early Archaean rocks is insensitive to metamorphic disturbances, so that ¹⁴⁶Sm–¹⁴²Nd chronometry can be reliably applied to a wide range of crustal rocks.

However, in view of the low initial abundance of ¹⁴⁶Sm, ¹⁴²Nd anomalies are expected to be extremely small (less than 30–40 p.p.m.), and their detection requires ultraprecise Nd isotope measurements. Until recently, reproducibility of Nd isotope ratios was limited to 15–20 p.p.m. (2σ) using the most precise instruments¹³. This precision barrier has severely hampered application of the ¹⁴⁶Sm–¹⁴²Nd chronometer to terrestrial differentiation because it precludes unambiguous resolution of small ¹⁴²Nd anomalies. The first claim by Harper and Jacobsen¹⁴ of a 33 p.p.m. excess in a 3.8-Gyr-old Isua metasediment was hence received with scepticism, especially since subsequent studies failed to detect any ¹⁴²Nd anomaly in the most ancient Archaean terranes^{15–17}. More recent

but preliminary analyses (less precise than this study) of two Amitsôq and one Isua sample of unknown location have also shown no anomaly¹⁸. However, in a thorough attempt to provide independent confirmation of Harper and Jacobsen's results, Sharma *et al.*¹³ found strong hints of a similar ^{142}Nd effect in two rocks from Isua, including the sample analysed by Harper and Jacobsen. Recent analyses by inductively coupled plasma mass spectrometry (ICP-MS) also hint at the presence of $^{142}\text{Nd}/^{144}\text{Nd}$ anomalies in several Isua metabasalts¹⁹. However, as emphasized by Sharma *et al.*¹³, these data lack the precision required to resolve with confidence such small $^{142}\text{Nd}/^{144}\text{Nd}$ anomalies. In order to constrain early Earth differentiation by the ^{146}Sm – ^{142}Nd chronometer, more precise analytical techniques are required.

Here we address this issue using TRITON, the new-generation thermal ionization mass spectrometer (TIMS). From an extensive Ames standard data set (see Methods), we estimate that a long-term reproducibility of 5 p.p.m. (2σ) can be achieved for $^{142}\text{Nd}/^{144}\text{Nd}$ (Fig. 1). This represents an improvement by a factor of 3–4 compared with previous TIMS analysis¹⁴, and allows quantitative resolution of ^{142}Nd anomalies as small as 10 p.p.m. in early Archaean rocks.

The seven analysed samples of amphibolite-facies metasediments are from the Isua greenstone belt (IGB) of West Greenland, and their sedimentary protoliths are interpreted as clastic erosion products from a mafic source region²⁰ (see Supplementary Information). Previous ^{147}Sm – ^{143}Nd studies⁶ of these 7 samples, together with 17 other closely related IGB metasediments, yielded a whole-rock regression age of $3,744 \pm 46$ Myr with initial $\varepsilon_{143}\text{Nd} = 1.9 \pm 0.6$.

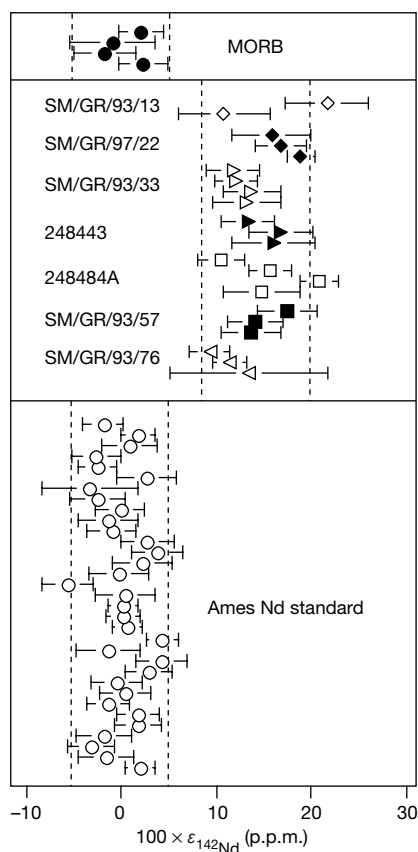


Figure 1 $\varepsilon_{142}\text{Nd}$ values of IGB metasediments compared with those of the Nd standard Ames ($N = 33$) and several replicates of a modern mid-ocean ridge basalt. Each Isua sample was measured 2, 3 or 4 times; these measurements were interspersed with several measurements of the Nd standard and the MORB sample. IGB metasediments display a well-defined excess averaging 15 p.p.m., whereas MORB data are indistinguishable from the Ames standard value.

Within error, this age agrees broadly with zircon U–Pb dates of $3,708 \pm 3$ Myr and $3,807 \pm 2$ Myr for nearby IGB felsic rocks²¹, demonstrating that the rocks remained closed systems for Sm–Nd since deposition. The regressed initial $\varepsilon_{143}\text{Nd}$ of 1.9 ± 0.6 therefore suggests that the basaltic source region of IGB metasediments is derived from a mantle reservoir with long-term depletion history.

The new ^{142}Nd measurements were obtained from May to August 2002 using the experimental procedure described in Methods. The results (Fig. 1) show that all IGB metasediment samples display well-resolved ^{142}Nd anomalies averaging $100 \times \varepsilon_{142}\text{Nd} = 15 \pm 4$ p.p.m. (2σ). In contrast, interspersed analysis of a modern mid-oceanic basalt yielded $^{142}\text{Nd}/^{144}\text{Nd}$ ratios identical to the Ames standard value within errors, thereby precluding experimental bias during rock sample analysis. No resolvable difference is obvious between individual samples at the 5 p.p.m. level of precision (2σ), suggesting isotopic homogeneity within the metasediment. This absence of significant variations of the $^{142}\text{Nd}/^{144}\text{Nd}$ ratio between analysed samples is consistent with the fairly homogeneous initial ^{143}Nd signature reported so far for IGB metasediments⁶. We therefore conclude that the observed ^{142}Nd anomalies reflect decay of now extinct ^{146}Sm in a Hadean silicate reservoir with an Sm/Nd ratio higher than chondritic.

One of our samples (SM/GR/93/13) was shown to carry a negative ^{182}W anomaly ($\varepsilon_{182}\text{W} = -1.23$), interpreted as reflecting the input of meteoritic tungsten to IGB metasediments at or after the so-called late heavy bombardment (LHB) at 3.8–4.0 Gyr ago²². It is therefore possible that a small fraction of Nd contained in IGB metasediments is extraterrestrial in origin. However, we regard it as highly implausible that observed ^{142}Nd anomalies reflect this extraterrestrial input. LHB meteoritic material is constrained to carry unradiogenic ^{182}W , so that observed ^{182}W anomalies are best explained by addition of chondrites or iron-meteorites²³. Clearly, iron meteorites can be ruled out immediately as a plausible cause of the ^{142}Nd anomaly because they do not contain significant amounts of rare-earth elements (REE). Furthermore, previous ^{142}Nd studies of whole-rock chondrites have shown no significant ^{142}Nd shift compared with terrestrial standards²⁴, so that input of chondritic neodymium is unlikely to produce a positive ^{142}Nd anomaly. Therefore, we believe that the observed ^{142}Nd excess is terrestrial in origin, and reflects Sm/Nd fractionation of the Earth's mantle at an early stage of its evolution.

We now examine the chronological implications of the new ^{142}Nd data using a simple two-stage model of isotopic evolution. In this model, the first stage represents evolution of a primitive chondritic mantle, and is followed by instantaneous differentiation of the crust–mantle system. The $^{142,143}\text{Nd}$ evolution of depleted mantle during the second stage is given by:

$$\left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}}\right)_t^{\text{DM}} = \left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}} + \left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}} (1 - e^{-\lambda_{147}t}) + \left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{DM}} (e^{-\lambda_{147}t_d} - e^{-\lambda_{147}t}) \quad (1)$$

$$\left(\frac{^{142}\text{Nd}}{^{144}\text{Nd}}\right)_t^{\text{DM}} = \left(\frac{^{142}\text{Nd}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}} - \frac{\left(\frac{^{146}\text{Sm}}{^{144}\text{Nd}}\right)_{t_0}^{\text{CHUR}}}{\left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}}} \times \left\{ \left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}} e^{-\lambda_{146}(t_0-t_d)} + \left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{DM}} [e^{-\lambda_{146}(t_0-t)} - e^{-\lambda_{146}(t_0-t_d)}] \right\} \quad (2)$$

where $t_0 = 4.566$ Gyr, t_d is the time of mantle differentiation, t_p is present-day time, t is any time after t_d , CHUR indicates chondritic

uniform reservoir, and DM indicates depleted mantle. $\lambda_{146} = 6.74 \text{ Gyr}^{-1}$ and $\lambda_{147} = 6.54 \times 10^{-3} \text{ Gyr}^{-1}$ are the respective decay constants of ^{146}Sm and ^{147}Sm .

Application of coupled $^{146,147}\text{Sm}$ chronometry to IGB metasediments with $\varepsilon_{143}\text{Nd} = 1.9 \pm 0.6$ and $100 \times \varepsilon_{142}\text{Nd} = 15 \pm 4 \text{ p.p.m.}$ yields a model age of differentiation of $4,460 \pm 115 \text{ Myr}$ with $^{147}\text{Sm}/^{144}\text{Nd} = 0.217 \pm 0.01$ for depleted mantle (Fig. 2). Because the observed $^{142,143}\text{Nd}$ signature may reflect partial remixing between depleted mantle and complementary crust (or undifferentiated mantle), the inferred $^{147}\text{Sm}/^{144}\text{Nd}$ is only a minimum estimate for the Hadean depleted mantle. In contrast, the $^{142,143}\text{Nd}$ model age has the property of remaining unchanged by such dilution processes, and is therefore a robust estimate.

It is often suggested that extraction of continental crust during the Hadean may have produced the light-REE depletion recorded in early Archaean rocks (see, for example, ref. 5). We now examine this hypothesis using the two-stage model described above. In this model, the $^{147}\text{Sm}/^{144}\text{Nd}$ ratio of depleted mantle complementary to modern-type continental crust is estimated to be 0.228–0.249 from the following equation:

$$\left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{DM}} = \left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}} + \frac{1}{\lambda_{147}T_{cc}} \left[\left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}}\right)_{t_p}^{\text{DM}} - \left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}}\right)_{t_p}^{\text{CHUR}} \right] \quad (3)$$

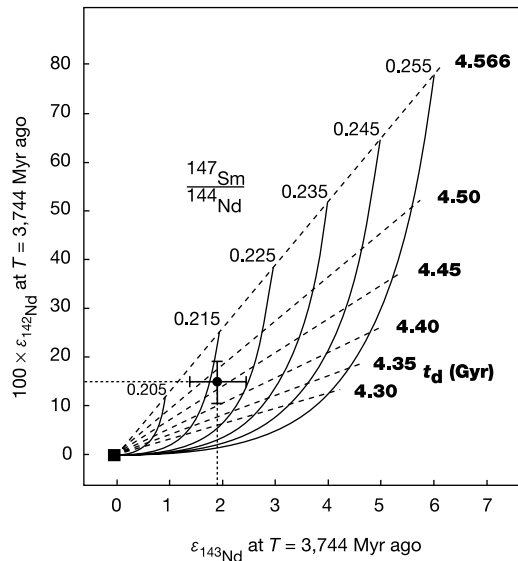


Figure 2 Snapshot of $^{142,143}\text{Nd}$ isotopic evolution of differentiated mantle at 3,744 Myr ago. The mantle first evolves as a closed-system chondritic reservoir and is then depleted at t_d . Dashed lines are loci of equal depletion ages, solid curves are loci of equal $^{147}\text{Sm}/^{144}\text{Nd}$ ratio of depleted mantle referenced to a present-day time. ε_{Nd} is defined as:

$$\varepsilon_{\text{Nd}}(t) = \left(\left(\frac{(^{142}\text{Nd}/^{144}\text{Nd})_t^{\text{DM}}}{(^{142}\text{Nd}/^{144}\text{Nd})_t^{\text{CHUR}}} - 1 \right) \times 10^4 \right)$$

^{142}Nd and ^{143}Nd isotopic signatures of depleted mantle are calculated using equations (1) and (2) with the $^{147}\text{Sm}/^{144}\text{Nd}$ ratio of depleted mantle ranging between 0.205 and 0.255 and ages of differentiation ranging between 4.3 Gyr and 4.566 Gyr. Other parameters used in the calculation are $(^{147}\text{Sm}/^{144}\text{Nd})_{t_0}^{\text{CHUR}} = 0.1966$, $(^{143}\text{Nd}/^{144}\text{Nd})_{t_0}^{\text{CHUR}} = 0.512638$, $(^{146}\text{Sm}/^{144}\text{Sm})_{t_0}^{\text{CHUR}} = 0.008$, $(^{142}\text{Nd}/^{144}\text{Nd})_{t_0}^{\text{CHUR}} = 1.141791$ and $(^{147}\text{Sm}/^{144}\text{Sm})_{t_0}^{\text{CHUR}} = 4.88899$. $(^{142}\text{Nd}/^{144}\text{Nd})_{t_0}^{\text{CHUR}}$ is assumed to be identical to the Ames standard value. Note that differentiation later than 4.3 Gyr ago does not result in a resolvable $\varepsilon_{142}\text{Nd}$ anomaly.

where $T_{cc} = 1.5\text{--}2.5 \text{ Gyr}$ is the mean age of present-day continental crust and $(^{143}\text{Nd}/^{144}\text{Nd})^{\text{DM}} = 0.51315$ is the average signature of mid-ocean-ridge basalts (MORBs). Note that similar results are obtained with simple mass-balance calculations, considering a Sm/Nd ratio of 0.19 for continental crust²³ and a complementary mantle representing 30% of total mantle mass (this corresponds to a 1.6% melt fraction for continental crust²⁴). Using this estimate, and constraining the mantle ^{142}Nd signature to $100 \times \varepsilon_{142}\text{Nd} = 15 \text{ p.p.m.}$ at 3,744 Myr ago, a model age of differentiation of 4.3–4.4 Gyr can be estimated from equation (2). The corresponding evolutions of $\varepsilon_{143}\text{Nd}$ and $\varepsilon_{142}\text{Nd}$ were calculated from equations (1) and (2) and are illustrated in Fig. 3. The results show that ^{143}Nd evolution generated with this model yields $\varepsilon_{143}\text{Nd}$ between 2.5 and 4 ε -units at 3,744 Myr ago, which appears significantly more radiogenic than observed in our Isua rocks. Conversely, if the ^{143}Nd signature of Archaean mantle is constrained to $\varepsilon_{143}\text{Nd} = 1.9$ at 3,744 Myr ago, then corresponding ^{142}Nd anomalies cannot exceed 5 p.p.m. at 3,744 Myr ago (Fig. 3). Therefore, it appears that early continental growth cannot fully satisfy the $^{142,143}\text{Nd}$ constraints for IGB metasediments. Hence, as suggested by several workers^{4,5}, differentiation of the Hadean mantle may have proceeded by way of mechanisms very different from those that prevailed during most of Earth's history. Because the model age inferred from coupled $^{146,147}\text{Sm}$ chronometry largely overlaps the time span of formation of the terrestrial planets, we believe that the observed ^{142}Nd anomaly most probably reflects differentiation of the Earth's mantle during ongoing terrestrial accretion.

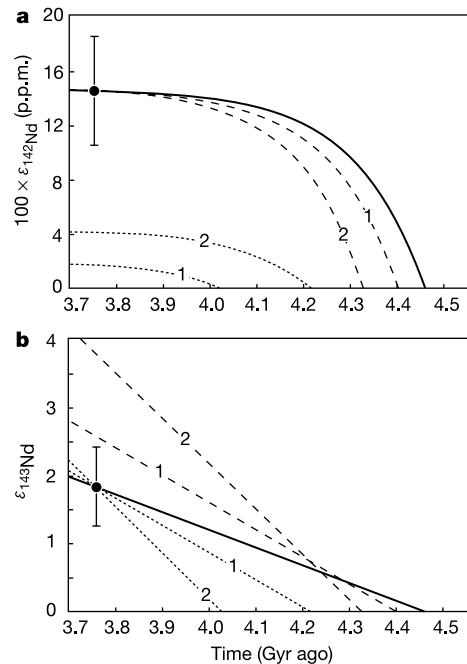


Figure 3 Isotopic evolution curves of depleted mantle following instantaneous differentiation of the mantle–crust system. **a**, $\varepsilon_{142}\text{Nd}$; **b**, $\varepsilon_{143}\text{Nd}$. Solid lines represent the two-stage model solution for the coupled $^{146,147}\text{Sm}$ systematics ($t_d = 4,460 \text{ Myr}$, $^{147}\text{Sm}/^{144}\text{Nd} = 0.217$). Dotted and dashed lines represent evolution curves of depleted mantle following extraction of modern-type continental crust. $^{147}\text{Sm}/^{144}\text{Nd}$ of depleted mantle is constrained to lie between 0.228 (curves labelled 1) and 0.249 (curves labelled 2) and were calculated using equation (3) with $T_{cc} = 2.5 \text{ Gyr}$ and 1.5 Gyr , respectively (see discussion in text). In the first model (dashed), the curves are calculated using equations (1) and (2) in order to fit $\varepsilon_{142}\text{Nd}$ in our Isua samples (shown as a solid circle in **a**) and a model age of mantle differentiation is derived ($t_d = 4.3\text{--}4.4 \text{ Gyr}$). Conversely, in the second model (dotted), isotopic evolution curves are calculated to fit $\varepsilon_{143}\text{Nd} = +1.9$ at 3,744 Myr (**b**). Note that neither of the two models yields satisfactory results, suggesting that extraction of modern-type continental crust cannot account for the $^{142,143}\text{Nd}$ signature of IGB metasediments.

Numerical simulations indicate that accretion of planetary material during the final 50–150 Myr of accretion proceeded by means of highly energetic collisions with intermediate-mass bodies (10^{26} – 10^{27} g)²⁵. The widespread existence of large ^{142}Nd anomalies (0.5–1 ϵ -units) in Nakhilite and Chassignite martian meteorites^{26,27} suggests that these large impacts are likely to induce large-scale mantle differentiation, either through a magma ocean stage or by impact-induced partial melting²⁸. If the Earth's mantle was partially or completely molten during the large-impact stage of accretion, the model age inferred from coupled $^{146,147}\text{Sm}$ chronometry would probably reflect its solidification and the end of terrestrial accretion. A protocrust formed at that time may have remained isolated from its mantle source until thermal convection began to operate. If the onset of mantle convection was delayed by several hundred million years, as suggested by recent numerical modelling²⁹, crustal recycling would have been relatively inefficient during the early Hadean, thereby favouring preservation of ^{142}Nd anomalies up to 3.75 Gyr ago. Alternatively, the observed ^{142}Nd excess may reflect the contribution of an older component to IGB metasediments—such as basaltic crust formed at a time when ^{142}Nd anomalies had not yet been erased by convective stirring. □

Methods

Long-term standard reproducibility

To estimate the external precision achievable with the new-generation TRITON TIMS, an extensive Ames standard data set was collected from March 2001 to August 2002. Data corrected for mass fractionation using the exponential law with $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$ yield a long-term reproducibility of 10 p.p.m. (2σ) for $^{142}\text{Nd}/^{144}\text{Nd}$ ($N = 33$). Although the exponential law is obviously the most appropriate to correct for Nd mass fractionation, corrected ratios display well-defined arrays when plotted in binary diagrams, which are largely responsible for their variability (see Supplementary Information). Similar correlations were recently described from Nd TIMS and ICP-MS data by Vance and Thirlwall³⁰. In agreement with these authors, we ascribe observed correlations between corrected isotope ratios to non-ideal mass fractionation during analysis. Although the origin of this mass fractionation bias remains uncertain, the clear correlations observed suggests that correction of this effect can be achieved provided that a second normalization ratio can be accurately measured. Hence, the following empirical law was parameterized to perform a second-order correction of our data for mass fractionation:

$$\left(\frac{R_{ij}^c}{R_j^c}\right) = \left(\frac{R_{ij}^e}{R_j^e}\right)^A \quad \text{with: } A = \frac{\ln(m_i/m_j)\ln(m_i/m_k)}{\ln(m_i/m_j)\ln(m_i/m_k)} \quad (4)$$

where the subscripts i, j, k and l stand for different Nd isotopes used in the calculations, R_{ij} represents the ratio of two isotopes of mass m_i and m_j , R^e and R^* represent respectively ratios corrected using the exponential law, and 'true' ratios corrected using equation (4). A first-order correction using the conventional exponential law with $R_{ij}^e = ^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$ yields R_{ij}^c . A second-order correction using equation (4) with $R_{ij}^e = ^{150}\text{Nd}/^{144}\text{Nd} = 0.236431$ yields R_{ij}^* . A is a mass-dependent coefficient that predicts well the slope defined by corrected data in a R_{ij}^c versus R_{ij}^e normalized plot.

The use of this method of correction substantially improved the external precision of Nd isotope ratios, with a reproducibility of 5 p.p.m. (2σ) for $^{142}\text{Nd}/^{144}\text{Nd}$. We emphasize that rock sample measurements display well-defined trends with slopes indistinguishable from those defined by the Ames standard data, thereby allowing application of equation (4) to our sample measurements. With the exception of the $^{142}\text{Nd}/^{144}\text{Nd}$ ratio (as well as $^{143}\text{Nd}/^{144}\text{Nd}$), for which the trends are systematically shifted toward higher values, all isotopic ratios are identical to the standard reference ratios within errors (see Supplementary Information).

Sample preparation and analysis

Sample aliquots of 50–200 mg were dissolved using HF + HNO₃ in Savillex Teflon beakers at ~140 °C. Separation of REE from the rock matrix was achieved using TRU-Spec resins. Cerium elimination was carried out using a liquid–liquid extraction technique³¹. After a second purification using AG50X8 cationic resin, Nd separation was done on a di(2-ethylhexyl)orthophosphoric acid (HDEHP)-coated Teflon chromatographic column. Total procedural blanks were 30–50 pg.

Nd (300–500 ng) was loaded on zone-refined Re filaments in 1 μl of 6N HCl, with H₃PO₄ subsequently added as an activator. Measurements were performed in the static mode, with ion beam intensities of $(4\text{--}6) \times 10^{-11}$ A. Analysis time was typically 6 h to ensure high internal precision ($2\sigma/n^{1/2} = 2\text{--}5$ p.p.m.). The presence of elements that could potentially form oxides or phosphates and interfere with Nd isotopes was carefully checked by performing mass scans using ion counting in the corresponding mass ranges and in the vicinity of Nd peaks. With the exception of alkali earths, which are in a mass range too low to produce oxides or phosphates around 140–150 a.m.u., no element was detected in sufficient amount to induce a significant effect on Nd isotope ratios. Ce and Sm isobaric interferences were also carefully monitored during rock sample measurements. With a few exceptions, $^{142}\text{Ce}/^{142}\text{Nd}$ was always <3 p.p.m., $^{144}\text{Sm}/^{144}\text{Nd} < 1$ p.p.m. and

$^{150}\text{Sm}/^{150}\text{Nd} < 10$ p.p.m., so that interference corrections were in most cases not required. When needed, ^{149}Sm and ^{140}Ce were monitored using ion counting in dynamic mode, and data corrected scan by scan. The non-radiogenic $^{148}\text{Nd}/^{144}\text{Nd}$ ratio was monitored carefully to detect atypical mass fractionation behaviour, isobaric interferences on normalizing isotopes (^{144}Nd and ^{146}Nd) and monitor instrument stability. For all rock sample measurements, $^{148}\text{Nd}/^{144}\text{Nd}$ was always identical to the Ames standard reference value within errors (see Supplementary Information).

Received 13 November 2002; accepted 9 April 2003; doi:10.1038/nature01668.

- Smith, A. D. & Ludden, J. N. Nd isotopic evolution of the Precambrian mantle. *Earth Planet. Sci. Lett.* **93**, 14–22 (1989).
- Moorbath, S., Whitehouse, M. J. & Kamber, B. S. Extreme Nd-isotope heterogeneity in the early Archaean—fact or fiction? Case histories from northern Canada and West Greenland. *Chem. Geol.* **135**, 213–231 (1997).
- Armstrong, R. L. Radiogenic isotopes: The case for crustal recycling on a near-steady-state no-continent-growth Earth. *Phil. Trans. R. Soc. Lond. A* **301**, 443–472 (1981).
- Chase, C. G. & Patchett, P. J. Stored mafic/ultramafic crust and early Archaean mantle depletion. *Earth Planet. Sci. Lett.* **91**, 66–72 (1988).
- Galer, S. J. G. & Goldstein, S. L. Early mantle differentiation and its thermal consequences. *Geochim. Cosmochim. Acta* **55**, 227–239 (1991).
- Kamber, B. S., Moorbath, S. & Whitehouse, M. J. Extreme Nd isotope heterogeneity in the early Archaean—fact or fiction? Case histories from northern Canada and West Greenland—Reply. *Chem. Geol.* **148**, 219–224 (1998).
- Allègre, C. J., Staudacher, T. & Sarda, P. Rare gas systematics: Formation of the atmosphere, evolution and structure of the Earth's mantle. *Earth Planet. Sci. Lett.* **81**, 127–150 (1986).
- Yin, Q. *et al.* A short timescale for terrestrial planet formation from Hf–W chronometry of meteorites. *Nature* **418**, 949–952 (2002).
- Kleine, T., Münker, C., Mezger, K. & Palme, H. Rapid accretion and early core formation on asteroids and the terrestrial planets from Hf–W chronometry. *Nature* **418**, 952–955 (2002).
- Schoenberg, R., Kamber, B. S., Collerson, K. D. & Eugster, O. New W-isotope evidence for rapid terrestrial accretion and early core formation. *Geochim. Cosmochim. Acta* **66**, 3151–3160 (2002).
- Lugmair, G. W., Shimamura, T., Lewis, R. S. & Anders, E. Samarium-146 in the early Solar System: Evidence from neodymium in the Allende meteorite. *Science* **222**, 1015–1018 (1983).
- Prinzhofer, D. A., Papanastassiou, D. A. & Wasserburg, G. J. Samarium-neodymium evolution of meteorites. *Geochim. Cosmochim. Acta* **56**, 797–815 (1992).
- Sharma, M., Papanastassiou, D. A., Wasserburg, G. J. & Dymek, R. F. The issue of the terrestrial record of ^{146}Sm . *Geochim. Cosmochim. Acta* **60**, 2037–2047 (1996).
- Harper, C. L. & Jacobsen, S. B. Evidence from coupled ^{147}Sm – ^{143}Nd and ^{146}Sm – ^{142}Nd systematics for very early (4.5 Gyr) differentiation of the Earth's mantle. *Nature* **360**, 728–732 (1992).
- Goldstein, S. L. & Galer, S. J. G. On the trail of early mantle differentiation: $^{142}\text{Nd}/^{144}\text{Nd}$ ratios of early Archaean rocks. *Eos* **73**, S323 (1992).
- McCulloch, M. T. & Bennett, V. C. Evolution of the early Earth: Constraints from ^{143}Nd – ^{142}Nd isotopic systematics. *Lithos* **30**, 234–255 (1993).
- Regelous, M. & Collerson, K. D. ^{147}Sm – ^{143}Nd , ^{146}Sm – ^{142}Nd systematics of early Archaean rocks and implications for crust–mantle evolution. *Geochim. Cosmochim. Acta* **60**, 3513–3520 (1996).
- Caro, G., Bourdon, B., Birck, J. L. & Moorbath, S. $^{142}\text{Nd}/^{144}\text{Nd}$ precise determination in early Archaean rocks. *Geochim. Cosmochim. Acta* **66**, A121 (2002).
- Boyet, M., Albarède, F., Télouk, P. & Rosing, M. ^{142}Nd anomaly confirmed at Isua. *Geochim. Cosmochim. Acta* **66**, A99 (2002).
- Boak, J. L. & Dymek, R. F. Metamorphism of the ca. 3800 Ma supracrustal rocks at Isua, West Greenland: Implications for early Archaean crustal evolution. *Earth Planet. Sci. Lett.* **59**, 155–176 (1982).
- Nuttman, A. P., Bennett, V. C., Friend, C. R. L. & Rosing, M. T. ≈ 3710 and ≥ 3790 Ma volcanic sequences in the Isua (Greenland) greenstone belt; structural and Nd isotope implications. *Chem. Geol.* **141**, 271–287 (1997).
- Schoenberg, R., Kamber, B. S., Collerson, K. D. & Moorbath, S. Tungsten isotope evidence from 3.8-Gyr metamorphosed sediments for early meteorite bombardment of the Earth. *Nature* **418**, 403–405 (2002).
- Goldstein, S. J. & Jacobsen, S. B. Nd and Sr isotopic systematics of river water suspended material: Implications for crustal evolution. *Earth Planet. Sci. Lett.* **87**, 249–265 (1988).
- Hofmann, A. W. Chemical differentiation of the Earth: The relationship between mantle, continental crust, and oceanic crust. *Earth Planet. Sci. Lett.* **90**, 297–314 (1988).
- Chambers, J. E. Making more terrestrial planets. *Icarus* **152**, 205–224 (2001).
- Harper, C. L., Nyquist, L. E., Bansal, B., Wiesmann, H. & Shih, C. Rapid accretion and early differentiation of Mars indicated by $^{142}\text{Nd}/^{144}\text{Nd}$ in SNC meteorites. *Science* **267**, 213–217 (1995).
- Borg, L. E., Nyquist, L. E., Taylor, L. A., Wiesmann, H. & Shih, C. Constraints on Martian differentiation processes from Rb–Sr and Sm–Nd isotopic analyses of the basaltic shergottite QUE 94201. *Geochim. Cosmochim. Acta* **61**, 4915–4931 (1997).
- Tonks, B. T. & Melosh, H. J. Magma ocean formation due to giant impacts. *J. Geophys. Res.* **98**, 5319–5333 (1993).
- Choblet, G. & Sotin, C. 3D thermal convection with variable viscosity: Can transient cooling be described by a quasi-static scaling law? *Phys. Earth Planet. Inter.* **119**, 321–336 (2000).
- Vance, D. & Thirlwall, M. An assessment of mass discrimination in MC-ICPMS using Nd isotopes. *Chem. Geol.* **185**, 227–240 (2002).
- Rehkämper, M., Gärtner, M., Galer, S. J. G. & Goldstein, S. L. Separation of Ce from other rare-earth elements with application to Sm–Nd and La–Ce chronometry. *Chem. Geol.* **129**, 201–208 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank C. Allègre, C. Jaupart and C. Langmuir for discussions, K. Burton for input at the earlier stage of this work, M. Delmotte for maintaining the TRITON in good condition, and M. Pierre for help with analytical chemistry. S.M. thanks P. Appel and the Isua Multidisciplinary Research Project for logistic and financial support of field work. This work was supported by the CNRS research programme 'Intérieur de la Terre'.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.C. (caro@ipgp.jussieu.fr).

Spontaneous emergence of leaders and followers in foraging pairs

Sean A. Rands^{*†}, Guy Cowlshaw[†], Richard A. Pettifor[†], J. Marcus Rowcliffe[†] & Rufus A. Johnstone^{*}

^{*} Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

[†] Institute of Zoology, Zoological Society of London, Regents Park, London NW1 4RY, UK

Animals that forage socially¹ often stand to gain from coordination of their behaviour^{2–5}. Yet it is not known how group members reach a consensus on the timing of foraging bouts. Here we demonstrate a simple process by which this may occur. We develop a state-dependent, dynamic game model⁶ of foraging by a pair of animals, in which each individual chooses between resting or foraging during a series of consecutive periods, so as to maximize its own individual chances of survival^{6,7}. We find that, if there is an advantage to foraging together^{1,2,8}, the equilibrium behaviour of both individuals becomes highly synchronized. As a result of this synchronization, differences in the energetic reserves of the two players spontaneously develop, leading them to adopt different behavioural roles. The individual with lower reserves emerges as the 'pace-maker' who determines when the pair should forage, providing a straightforward resolution to the problem of group coordination. Moreover, the strategy that gives rise to this behaviour can be implemented by a simple 'rule of thumb' that requires no detailed knowledge of the state of other individuals.

When animals forage, they have to make many decisions^{7,10}: when to rest or to forage, when to leave a patch, and which food items to seek. Their behaviour is likely to reflect a trade-off between the risks of starvation and predation^{11,12}, changing over the short-term in response to both their energy reserves and the environment⁶. For animals living in social groups, foraging is further complicated by the actions of other individuals¹: it will usually be safer to forage at the same time as other members of the group^{2,8} (although this may also entail competition for food²), favouring some coordination of activity^{3–5,13}. How can group members reach a consensus on the timing of foraging bouts? Models of self-organization in biological systems⁹ have shown how coordinated patterns of activity can emerge spontaneously in groups of individuals following simple, predetermined behavioural rules. Here we describe a dynamic, game-theoretical model of foraging in a pair (see Supplementary Information for details), in which the evolutionarily stable rule governing the timing of individual foraging bouts leads in this way to a simple resolution of the problem of group coordination.

Our model considers decisions of the pair of animals over a large number of discrete time periods. During a given period, each individual can choose to forage or to rest. Foraging increases energy reserves, and hence decreases the forager's starvation risk. However,

it may also entail an increased risk of predation^{11,12}. Previously, dynamic programming^{6,14,15} has been successfully used to examine the effects of trade-offs between starvation and predation risk on the short-term, minute-to-minute behaviour of individual foragers. Dynamic games use similar techniques⁶, but allow us to model the behaviour of groups. Although state-dependent models of group-foraging have been developed previously, looking at optimal group size^{16,17}, diurnal routines¹⁸, foraging strategy¹, sentinel behaviour¹⁹ and social hierarchy effects^{20,21}, we believe the model presented here to be the first that has explicitly considered the coordination of foraging among individuals within a group. In the form presented here, the model only considers a pair of players. The same approach could in principle be used to generate predictions about the behaviour of foragers in larger groups, but the computations involved have proved to be unfeasibly time-consuming.

The state of both players is represented by their level of energetic reserves, bounded by a maximum above which a player cannot gain further reserves and a minimum level at which the player is assumed to have starved. We assume that each player has exact knowledge of both its own and its partner's state (though, as we discuss later, this assumption ultimately proves unnecessary). At the beginning of a period, each player decides whether to rest or to forage. If a player rests, on average it incurs a net loss in reserves. If a player forages, on average it receives a net gain. Both energetic costs and the amount of food found during a period vary probabilistically with a known distribution. All behaviours have a risk of predation attached to them^{22,23}—foraging alone is assumed to be more (or equally) risky than foraging in a pair, which itself is more (or equally) risky than resting. The results presented in the figures are based on the assumption that the average amount of energy gained by a foraging individual is independent of the number of others foraging at the same time, although we also explored the effects of relaxing this assumption.

We calculated the evolutionarily stable strategy (ESS) for players in this game, for a range of different parameter values. In particular, we focused on the effect of omitting or incorporating a benefit of foraging together—we obtained closely similar results when foraging together gave either a safety benefit (that is, reduced predation risk) or an energetic benefit (that is, higher mean individual net gain) compared to foraging alone. The ESS in any one case was found using an iterated, damped best-response procedure^{6,24}, incorporating some likelihood of errors in decision making²⁵. We were then able to predict the long-term stable state distribution, and the expected behaviour of a pair of individuals that adopt this strategy. In particular, we examined the expected level of behavioural coordination at equilibrium, quantified using the relative disequilibrium parameter²⁶ D' (where $D' = 1$ implies that pair members always adopt the same behaviour, $D' = 0$ implies that their actions are uncorrelated, and $D' = -1$ implies that they always adopt different behaviours). We also examined the mean state of players, the mean difference in state between them, and the correlation between their states.

When there is no benefit to foraging in a pair, both players behave independently, that is, the action adopted by either is not affected by the other's state. The optimal policy under these conditions, as in optimization models of individual behaviour⁶, is to forage when an individual's own state drops below some threshold value, and otherwise to rest (Fig. 1a). As a result, the energetic reserves of both players settle at similar levels, fluctuating around the threshold value (because a player below the threshold will forage until its reserves rise above the cut-off point, after which it should rest until its reserves again fall below this value). This gives a single peak in the stable state distribution for the pair (Fig. 1b). In this simple case, there is no correlation between the behaviours of the players ($D' = 0$).

By contrast, when we introduce a benefit to foraging together, through a lower predation risk than that incurred when foraging alone (or through an increase in mean energetic gain, although this

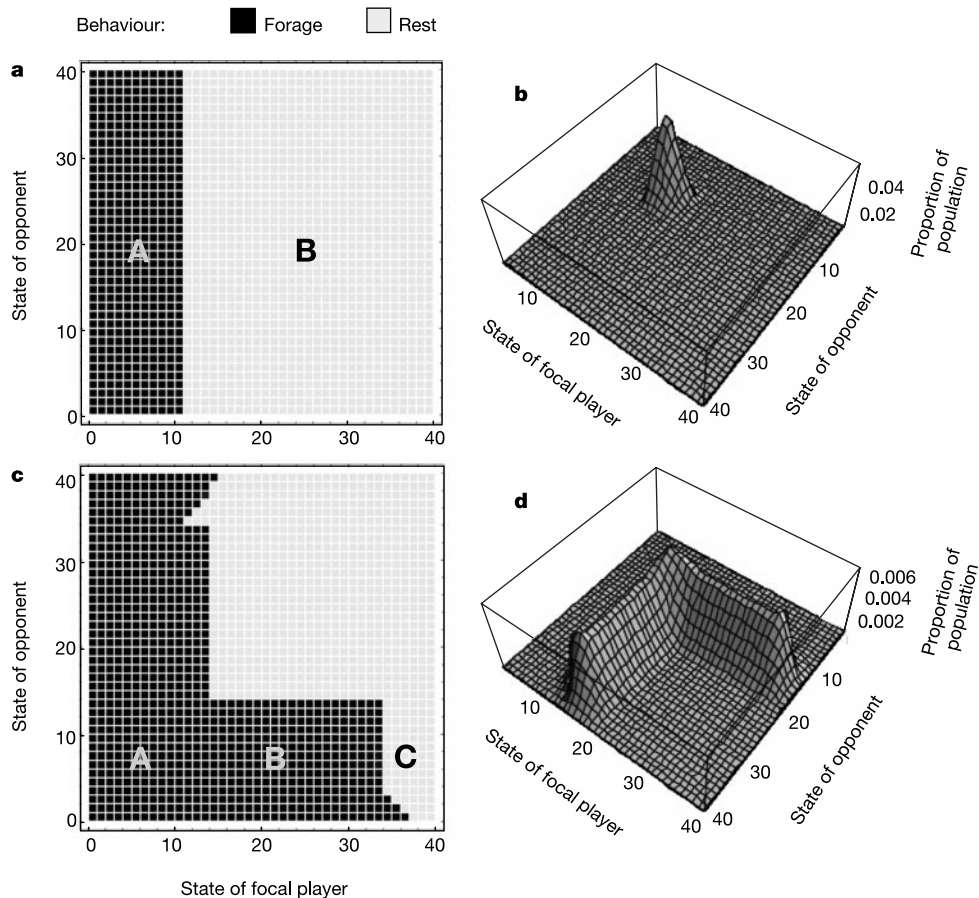


Figure 1 Evolutionarily stable policies (left column) and the probabilities that a pair of individuals following the policy exhibit any particular combination of states (right column), with differing benefits of foraging together. In **a** and **b**, there is no benefit to foraging together (following the standard parameter set given in Supplementary Information, but with $m_{FT} = m_{FA} = 0.0005$, where m_{FT} and m_{FA} are respectively the predation risks of foraging together and foraging alone). Policy **a** shows that if the state of the focal player falls below a threshold value, it should forage (zone A), whereas if its state rises above this threshold, it should rest (zone B). Both players follow the same policy, and so the states of both remain near the threshold (**b**). In **c** and **d**, foraging together yields a safety benefit (following the standard parameter set). Policy **c** has a similar focal state threshold to policy **a**, below which the focal player must forage to avoid starvation (zone A). However, there is

also a region above the threshold (B) in which it should forage—this occurs when the partner has a low state, and so is likely to forage. The focal player therefore also forages (despite having high reserves) because it incurs a lower predation risk if it does so, rather than waiting until its reserves have fallen to a critical level and then potentially having to forage alone (incurring a greater predation risk). This is true for all but the highest of the focal player's states (zone C)—here, the player should always rest, because the ceiling on the level of reserves the animal can store negates any extra benefits of foraging. The state distribution (**d**) for this policy demonstrates that if one player has reserves above the threshold, its colleague should be on the threshold. Hence, one player always has low reserves when there is some benefit to foraging together.

latter possibility is not illustrated), the equilibrium behaviour of each individual becomes highly dependent upon the reserves of its partner, as well as its own state. Figure 1c describes a typical policy when there is a safety benefit of foraging together. Each individual forages when either its own state or that of the other player drops below some threshold value (unless the individual making the decision is approaching maximum reserve levels, in which case nothing further can be gained by foraging). Sensitivity analyses show that this result is extremely robust, as systematic and thorough variation of the parameters defined (such as the mean energetic gains, costs or predation risk of the behavioural options) did not lead to any qualitative differences between the strategies that were calculated.

The above strategy is beneficial because an individual that forages when its partner's reserves are low is likely to gain the safety benefit of joint activity (as under these circumstances the partner will probably also choose to forage). The correlation between the behaviour of the players following the policy shown in Fig. 1c is $D' = 0.9985$, which indicates that their activities are very tightly synchronized indeed. Surprisingly, however, the equilibrium correlation between the states of the players is negative ($\rho = -0.589$), implying that if one player has low reserves, its partner is likely to

have high reserves. This negative correlation in state is a consequence of behavioural synchronization. When foraging behaviour becomes tightly coupled, there is no opportunity for a player with lower reserves to forage while its partner rests, and so to reduce the difference in state between them. Consequently, the mean equilibrium difference in state between the two players builds up to 13.08 ± 7.91 state units (Fig. 1d); this value is approximately ten times greater than the mean difference of 1.77 ± 1.42 state units, which occurs when there is no benefit to foraging together and behaviour is unsynchronized (Fig. 1b).

This separation of the players into 'fat' (high reserves) and 'lean' (low reserves) roles persists for a much longer period of time than if there is no advantage to foraging together (Fig. 2). The players occasionally switch roles when, through a run of bad luck, the reserves of the 'fat' player drop below those of the 'lean' player, but this occurs only rarely: the median time for which one individual retains the 'fat' role is 109.07 periods (approximately 50 times longer than the equivalent median time of 2.28 when there is no advantage to foraging together). In other words, the model predicts that a temporarily stable asymmetry in state will emerge between two otherwise identical players that are both following an identical set of rules.

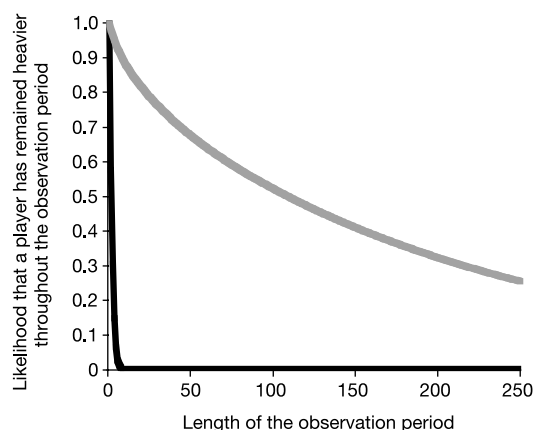


Figure 2 Likelihood that 'fat' and 'lean' roles persist over time. Given that one individual has higher energy reserves than its partner at a known point in time, this graph shows the probability that it will still have higher energy reserves (that is, that no switch in roles will have occurred) after any given length of time, assuming that both players follow the equilibrium strategy. The black line shows the results of the model when foraging together yields no benefit (and behaviour is unsynchronized); the grey line shows results when foraging together yields a safety advantage—see Fig. 1 for parameters.

The difference in state between the two players leads to a simple resolution of the coordination problem. The reserves of the 'lean' player (as illustrated in Fig. 1d) settle around the lower critical state threshold, at which its behaviour is dictated solely by its own energy reserves, regardless of the state of its partner. By contrast, the reserves of the 'fat' player rise above this threshold, and it will therefore respond to the state (and probable behaviour) of the 'lean' player. As a result, the activities of both members of the pair are dictated by the player with the lower reserves, which becomes the 'pace-maker' for the group. Both players will forage when the pace-maker's reserves drop below the lower critical threshold; both will rest when the pace-maker's reserves rise above this threshold.

We have also examined the case in which foraging together is disadvantageous (owing, for instance, to greater predation risk than when foraging alone). This generates policies that lead to anti-synchronous behaviour, with a correlation approaching $D' = 1$. However, this finding is unlikely to prove biologically relevant. If social foraging has detrimental effects on the fitness of an individual, then foraging groups are unlikely to form in the first place, and the problem of behavioural coordination does not arise.

To sum up, our results indicate that the problem of group coordination may be easily resolved through the spontaneous emergence of temporary 'leaders' and 'followers', owing to the build-up of differences in energetic state (even if group members experience identical chances of foraging success and predation). In support of this prediction, studies of foraging in groups of fish have shown that leadership decisions may often be made by individuals with lower reserves^{27–30}. Furthermore, we have shown that such coordination is possible even when individuals follow an evolutionarily stable strategy that maximizes their own chances of survival irrespective of the risks to other group members (in the same way that coordinated patterns of sentinel behaviour may emerge within groups of 'selfish' individuals¹⁹). This outcome is not the result of a communal group decision⁵, but emerges from the interaction between individuals who each make their own choices about when to forage or rest; as in other models of self-organization, the pattern of group activity is the product of individual decisions⁹.

We emphasize that although the focal player's decisions in the model are based on an exact knowledge of the energetic reserves of both itself and its partner, this information is not necessary for implementation of the ESS policy that we have derived. Instead, the ESS policy can be closely approximated by a much simpler rule of

thumb: 'I should forage if either my reserves have fallen below a certain threshold value, or my partner chooses to forage'. This tactic requires only the ability to observe and react (quickly) to a change in the partner's behaviour, rather than detailed knowledge of its state. Nevertheless, it yields identical behaviour to the calculated ESS policy, because the focal individual's partner will forage whenever its reserves drop below the threshold value. The simplicity of such a rule of thumb suggests that it may easily evolve, and implies that the solution we propose to the problem of coordination does not require any exceptional cognitive abilities. It also shows that very simple individual rules can be sufficient to generate seemingly complex patterns of social behaviour⁹.

Received 30 January; accepted 28 March 2003; doi:10.1038/nature01630.

- Giraldeau, L.-A. & Caraco, T. *Social Foraging Theory* (Princeton Univ. Press, Princeton, 2000).
- Krause, J. & Ruxton, G. D. *Living in Groups* (Oxford Univ. Press, Oxford, 2002).
- Conradt, L. Could asynchrony in activity between the sexes cause intersexual social segregation in ruminants? *Proc. R. Soc. Lond. B* **265**, 1359–1363 (1998).
- Ruckstuhl, K. E. & Neuhaus, P. Sexual segregation in ungulates: a comparative test of three hypotheses. *Biol. Rev.* **77**, 77–96 (2002).
- Conradt, L. & Roper, T. J. Group decision-making in animals. *Nature* **421**, 155–158 (2003).
- Houston, A. I. & McNamara, J. M. *Models of Adaptive Behaviour: an Approach Based on State* (Cambridge Univ. Press, Cambridge, 1999).
- Cuthill, I. C. & Houston, A. I. in *Behavioural Ecology: an Evolutionary Approach*, 4th edn (eds Krebs, J. R. & Davies, N. B.) 97–120 (Blackwell Science, Oxford, 1997).
- Hamilton, W. D. Geometry for the selfish herd. *J. Theor. Biol.* **31**, 295–311 (1971).
- Camazine, S. et al. *Self-organization in Biological Systems* (Princeton Univ. Press, Princeton, 2001).
- Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton Univ. Press, Princeton, 1986).
- Lima, S. L. Predation risk and unpredictable feeding conditions: determinants of body mass in birds. *Ecology* **67**, 377–385 (1986).
- McNamara, J. M. & Houston, A. I. The value of fat reserves and the tradeoff between starvation and predation. *Acta Biotheor.* **38**, 37–61 (1990).
- Rook, A. J. & Penning, R. D. Synchronisation of eating, ruminating and idling activity by grazing sheep. *Appl. Anim. Behav. Sci.* **32**, 157–166 (1991).
- Mangel, M. & Clark, C. W. *Dynamic Modeling in Behavioral Ecology* (Princeton Univ. Press, Princeton, 1988).
- Houston, A. I., Clark, C. W., McNamara, J. M. & Mangel, M. Dynamic models in behavioural and evolutionary ecology. *Nature* **332**, 29–34 (1988).
- Clark, C. W. The lazy, adaptable lions: a Markovian model of group foraging. *Anim. Behav.* **35**, 361–368 (1987).
- Mangel, M. Resource divisibility, predation and group formation. *Anim. Behav.* **39**, 1163–1172 (1990).
- Székely, T., Sozou, P. D. & Houston, A. I. Flocking behaviour of passerines: a dynamic model for the non-reproductive season. *Behav. Ecol. Sociobiol.* **28**, 203–213 (1991).
- Bednekoff, P. A. Mutualism among safe, selfish sentinels: a dynamic game. *Am. Nat.* **150**, 373–392 (1997).
- Clark, C. W. & Ekman, J. Dominant and subordinate fattening strategies: a dynamic game. *Oikos* **72**, 205–212 (1995).
- Richards, S. A. Temporal partitioning and aggression among foragers: modeling the effects of stochasticity and individual state. *Behav. Ecol.* **13**, 427–438 (2002).
- Lima, S. L. Stress and decision making under the risk of predation: recent developments from behavioral, reproductive, and ecological perspectives. *Adv. Stud. Behav.* **27**, 215–290 (1998).
- Lima, S. L. & Dill, L. M. Behavioral decisions made under the risk of predation: a review and prospectus. *Can. J. Zool.* **68**, 619–640 (1990).
- Enquist, M. & Leimar, O. Evolution of fighting behavior: decision rules and assessment of relative strength. *J. Theor. Biol.* **102**, 387–410 (1983).
- McNamara, J. M., Webb, J. N., Collins, E. J., Székely, T. & Houston, A. I. A general technique for computing evolutionarily stable strategies based on errors in decision-making. *J. Theor. Biol.* **189**, 211–225 (1997).
- Lewontin, R. C. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **49**, 49–67 (1964).
- Gotceitas, V. & Godin, J.-G. J. Foraging under the risk of predation in juvenile Atlantic salmon (*Salmo salar* L.): effects of social status and hunger. *Behav. Ecol. Sociobiol.* **29**, 255–261 (1991).
- Krause, J., Bumann, D. & Todt, D. Relationship between the position preference and nutritional state of individuals in schools of juvenile roach (*Rutilus rutilus*). *Behav. Ecol. Sociobiol.* **30**, 177–180 (1992).
- Krause, J. The relationship between foraging and shoal position in a mixed shoal of roach (*Rutilus rutilus*) and chub (*Leuciscus cephalus*): a field study. *Oecologia* **93**, 356–359 (1993).
- Krause, J., Reeves, P. & Hoare, D. Positioning behaviour in roach shoals: the role of body length and nutritional state. *Behaviour* **135**, 1031–1039 (1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank F. Devas, A. Randall, K. E. Ruckstuhl and S. R. X. Dall for discussions. This work was supported by the Natural Environment Research Council.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.A.R. (s.rands@zoo.cam.ac.uk).

GSK-3 α regulates production of Alzheimer's disease amyloid- β peptides

Christopher J. Phiel^{*†}, Christina A. Wilson^{†‡}, Virginia M.-Y. Lee[‡] & Peter S. Klein^{*}

^{*} Department of Medicine, Division of Hematology–Oncology and Howard Hughes Medical Institute, and [‡] Center for Neurodegenerative Disease Research, Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, 415 Curie Blvd, Philadelphia, Pennsylvania 19104-6148, USA [†] These authors contributed equally to this work

Alzheimer's disease is associated with increased production and aggregation of amyloid- β (A β) peptides¹. A β peptides are derived from the amyloid precursor protein (APP) by sequential proteolysis, catalysed by the aspartyl protease BACE², followed by presenilin-dependent γ -secretase cleavage³. Presenilin interacts with nicastrin^{4,5}, APH-1 and PEN-2 (ref. 6), all of which are required for γ -secretase function. Presenilins also interact with α -catenin, β -catenin^{7,8} and glycogen synthase kinase-3 β (GSK-3 β)^{9–11}, but a functional role for these proteins in γ -secretase activity has not been established. Here we show that therapeutic concentrations of lithium, a GSK-3 inhibitor¹², block the production of A β peptides by interfering with APP cleavage at the γ -secretase step, but do not inhibit Notch processing. Importantly, lithium also blocks the accumulation of A β peptides in the brains of mice that overproduce APP. The target of lithium in this setting is GSK-3 α , which is required for maximal processing of APP. Since GSK-3 also phosphorylates tau protein, the principal component of neurofibrillary tangles, inhibition of GSK-3 α offers a new approach to reduce the formation of both amyloid plaques and neurofibrillary tangles, two pathological hallmarks of Alzheimer's disease.

Amyloid plaques are composed primarily of 40- and 42-amino-acid peptides—A β ₄₀ and A β ₄₂, respectively—derived from APP¹. To investigate the role of GSK-3 in the production of A β ₄₀ and A β ₄₂, we treated Chinese hamster ovary cells that stably expressed APP (CHO-APP₆₉₅ cells) with lithium chloride, a direct inhibitor of GSK-3 α and GSK-3 β ¹³, for 24 h, and then measured A β ₄₀ and A β ₄₂ secretion by sandwich ELISA (enzyme-linked immunosorbent assay). LiCl inhibited the production of both A β ₄₀ and A β ₄₂ with an IC₅₀ (half-maximal inhibitory concentration) of 1–2 mM—within the therapeutic range for lithium—whereas sodium chloride had no effect (Fig. 1a), consistent with a recent report using transient overexpression of the APP carboxy terminus C100 in COS7 cells¹⁴. The reduction of A β was confirmed by metabolic labelling of CHO-APP₆₉₅ cells with ³⁵S-methionine in the presence or absence of LiCl. A β peptides could be immunoprecipitated from the medium of control cells, but treatment with LiCl or the γ -secretase inhibitor DFK-167 (ref. 15) markedly reduced the levels of secreted A β peptides (Fig. 1b).

Lithium did not affect steady-state levels of APP (Fig. 1c) or the levels of N- and C-terminal fragments of presenilin 1 (PS1), and it did not interfere with the detection of A β peptides (data not shown). Lithium therefore appeared to reduce the level of A β peptides at a post-translational step, such as APP processing or A β stability. To test this, we measured the accumulation of APP processing intermediates in the presence of LiCl. α -Secretase or β -secretase cleavage of APP generates APP C-terminal fragments (CTFs), which are then cleaved by γ -secretase. If γ -secretase is inhibited—for example, with DFK-167—then APP CTFs accumulate¹⁶. Exposure of CHO-APP₆₉₅ cells to 5 mM lithium caused the accumulation of APP CTFs (Fig. 1c), similar

to treatment with DFK-167. The accumulation of the APP CTFs shows that lithium inhibits APP processing at or before the γ -secretase cleavage step.

γ -Secretase is also required for the release of the Notch intracellular domain (NICD)¹⁵. However, lithium did not inhibit Notch processing, as assessed by immunoblotting with an antibody specific for cleaved Notch (Fig. 1d). Δ E-V1744K, a mutated form of Notch that is not cleaved by γ -secretase¹⁷, was not detected with this antibody (Fig. 1d, lane 3). Thus, lithium is not a general inhibitor of γ -secretase under these conditions, and might

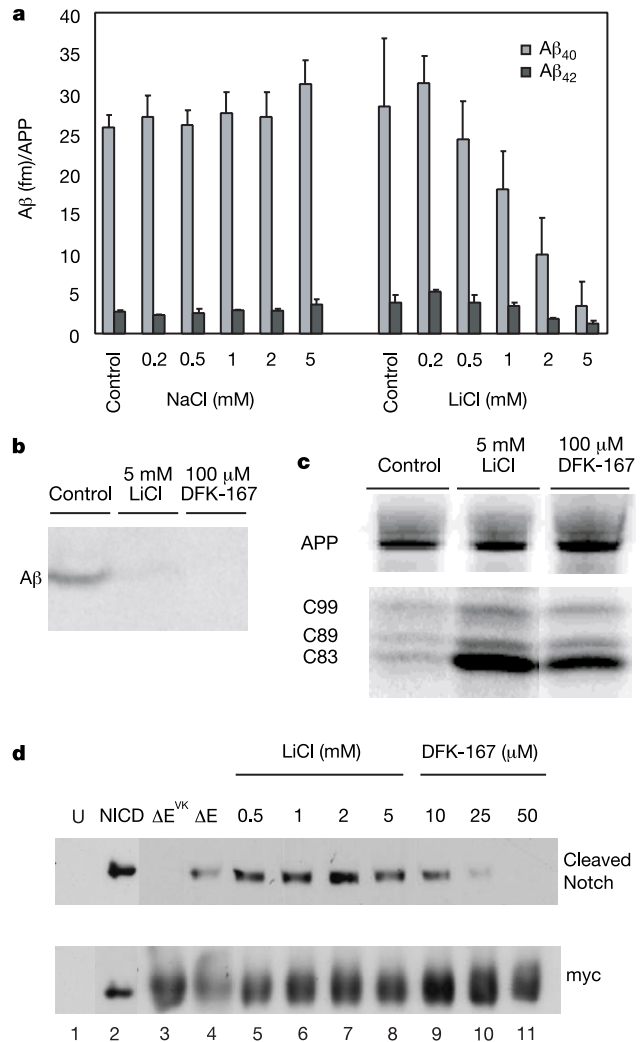


Figure 1 Lithium reduces secreted A β ₄₀ and A β ₄₂, and causes accumulation of APP C-terminal fragments. **a**, CHO-APP₆₉₅ cells were treated with NaCl or LiCl for 24 h. A β ₄₀ (light bars) and A β ₄₂ (dark bars) levels (in femtomoles) were measured by ELISA and normalized to intracellular full-length APP (determined by quantitative immunoblots). Error bars represent standard deviation. **b**, Immunoprecipitated A β peptides secreted from CHO-APP₆₉₅ cells after 24 h in culture with either 5 mM LiCl (lane 2) or 100 μ M DFK-167 (lane 3). **c**, CHO-APP₆₉₅ cells were treated with 5 mM LiCl or 100 μ M DFK-167 for 24 h. Intracellular APP and C-terminal fragments (C99, C89 and C83) were immunoprecipitated, separated by SDS–PAGE, and visualized by PhosphorImager. **d**, CHO-APP₆₉₅ cells were transfected with Δ E-Notch, which is constitutively cleaved by γ -secretase¹⁷ (lanes 4–11), treated with LiCl or DFK-167 for 24 h, and analysed by immunoblotting with an antibody to processed Notch. The NICD is a positive control for the antibody (lane 2); the Δ E^{WK} form of Notch is not cleaved by γ -secretase (lane 3). All Notch constructs have C-terminal myc tags. Overexpression of GSK-3 α or GSK-3 β had no effect on Notch cleavage (data not shown). U, untransfected.

specifically affect γ -secretase activity towards APP, or alter access of APP to the γ -secretase complex.

Although lithium is a direct, selective inhibitor of GSK-3 (ref. 13), it also inhibits inositol monophosphatase, related phosphomonoesterases, and phosphoglucomutase¹². To test the hypothesis that

lithium reduces A β production through inhibition of GSK-3, we treated CHO-APP₆₉₅ cells with the structurally unrelated GSK-3 inhibitor kenpaullone¹⁸. Kenpaullone markedly reduced both A β_{40} and A β_{42} (Fig. 2a), with 50% reduction at 2 μ M and >90% reduction at 5 μ M. Kenpaullone also inhibits cyclin-dependent kinases (cdks); although this action requires at least 20-fold higher concentrations *in vitro* than are needed for the inhibition of GSK-3 (ref. 18), we nevertheless examined the effect of roscovitine, a cdk inhibitor that does not inhibit GSK-3 (ref. 18), on CHO-APP₆₉₅ cells. Roscovitine had no effect on A β_{40} and A β_{42} (Fig. 2a), showing that kenpaullone does not inhibit A β production through the inhibition of cdks. Both kenpaullone and lithium inhibited GSK-3 in these cells, as indicated by accumulation of β -catenin protein (Fig. 2b) and activation of a β -catenin/Tcf-dependent reporter (Fig. 2c).

PS1 also interacts with β -catenin, and regulates β -catenin protein levels and subcellular localization^{10,11}. Because GSK-3 inhibitors cause accumulation of β -catenin¹², it is possible that β -catenin has a role as a downstream effector of PS1 or as a direct modulator of PS1 function. Therefore, we tested whether increasing the levels of β -catenin protein could mimic the effect of lithium and kenpaullone on A β production. Overexpression of β -catenin did not affect A β production (Fig. 2d) under conditions that activated the β -catenin-dependent reporter (data not shown); infection with semliki forest virus (SFV) encoding β -catenin also had no effect on A β production (data not shown). These observations show that an elevated level of β -catenin is not sufficient to cause the decrease in A β seen with GSK-3 inhibitors.

These data show that two structurally unrelated inhibitors of GSK-3 reduce the production of A β peptides. Although the most parsimonious explanation is that these inhibitors act through the inhibition of GSK-3, the possibility remains that these two agents fortuitously inhibit distinct, as yet unknown targets involved in APP processing. To confirm that GSK-3 regulates APP processing, we reduced the expression of endogenous GSK-3 using RNA interference (RNAi)¹⁹. In CHO-APP₆₉₅ cells, transfection of short interfering RNAs (siRNAs) directed against GSK-3 α and GSK-3 β reduced GSK-3 protein levels in an isoform-specific manner, whereas a control siRNA had no effect on GSK-3 expression (Fig. 3a). Selective reduction of GSK-3 α protein expression decreased A β_{40} and A β_{42} levels by 45% and 43%, respectively (Fig. 3b). These data show that GSK-3 α is required for maximal production of A β_{40} and A β_{42} , and suggest that GSK-3 β can antagonize APP processing in this setting. By contrast, reduction of GSK-3 β protein levels resulted in a modest increase in A β_{40} and A β_{42} concentration; this is surprising, as GSK-3 α and GSK-3 β are 97% identical within the kinase domains and are apparently redundant in the Wnt pathway, although not in the regulation of NF- κ B (nuclear factor- κ B)²⁰. However, the amino- and carboxy-terminal sequences of GSK-3 α and GSK-3 β are dissimilar, and this divergence may account for functional differences in the regulation of APP processing.

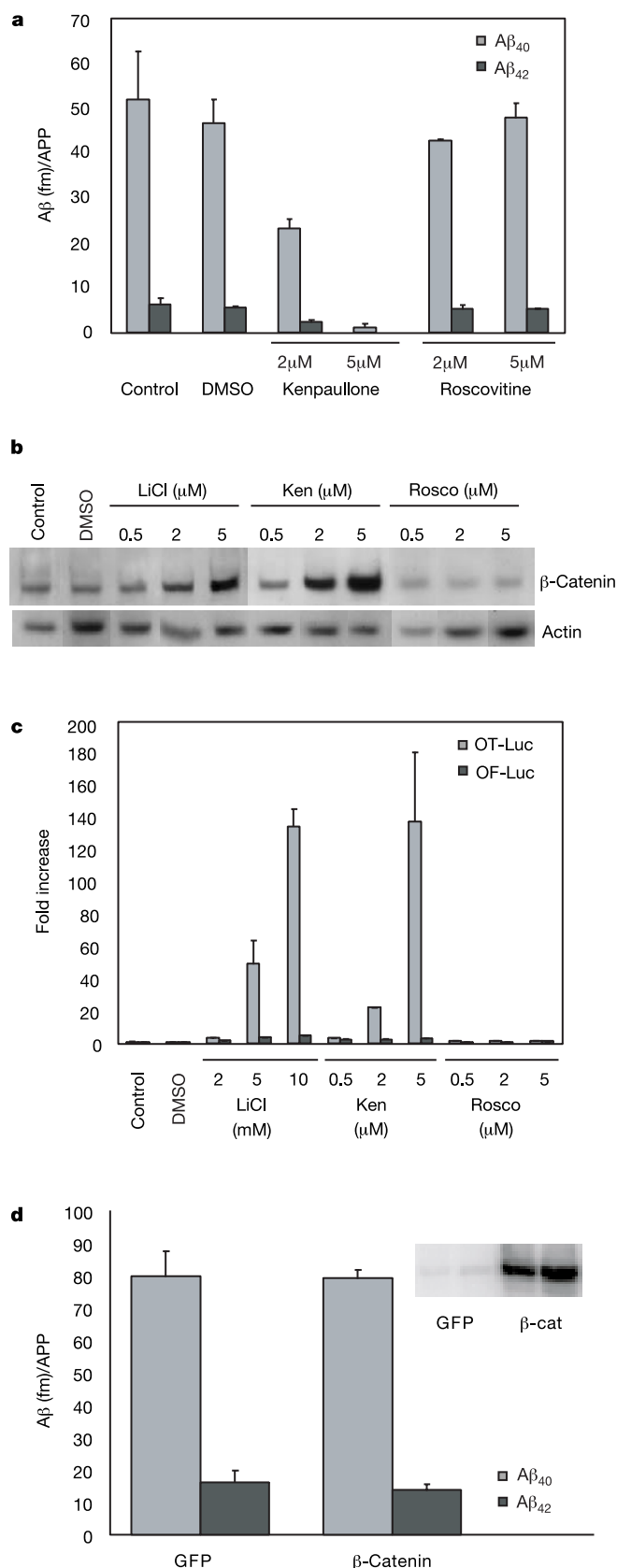


Figure 2 GSK-3 inhibitors reduce A β independently of β -catenin stabilization.

a, CHO-APP₆₉₅ cells were treated with kenpaullone or roscovitine (cdk inhibitor) for 24 h. Secreted A β_{40} and A β_{42} were measured as in Fig. 1. **b**, β -Catenin levels were assessed by immunoblot in CHO-APP₆₉₅ cells treated with kenpaullone, lithium or roscovitine. Immunoblot for actin is a loading control. Control, untreated cells; DMSO, dimethylsulphoxide; Ken, kenpaullone; Rosco, roscovitine. **c**, CHO-APP₆₉₅ cells transfected with the β -catenin-responsive reporter OT-luciferase (OT-Luc) or the mutated reporter OF-luciferase (OF-Luc) were treated with LiCl, kenpaullone or roscovitine for 24 h, and luciferase activity was measured. **d**, CHO-APP₆₉₅ cells were transfected with green fluorescent protein (GFP; control) or β -catenin. Secreted A β_{40} and A β_{42} levels were assessed as in Fig. 1. Inset shows immunoblot of endogenous (GFP) and overexpressed (β -cat) β -catenin in duplicate lanes.

Because either lithium or GSK-3 α depletion reduces A β levels, we predicted that raising GSK-3 α levels would enhance A β production.

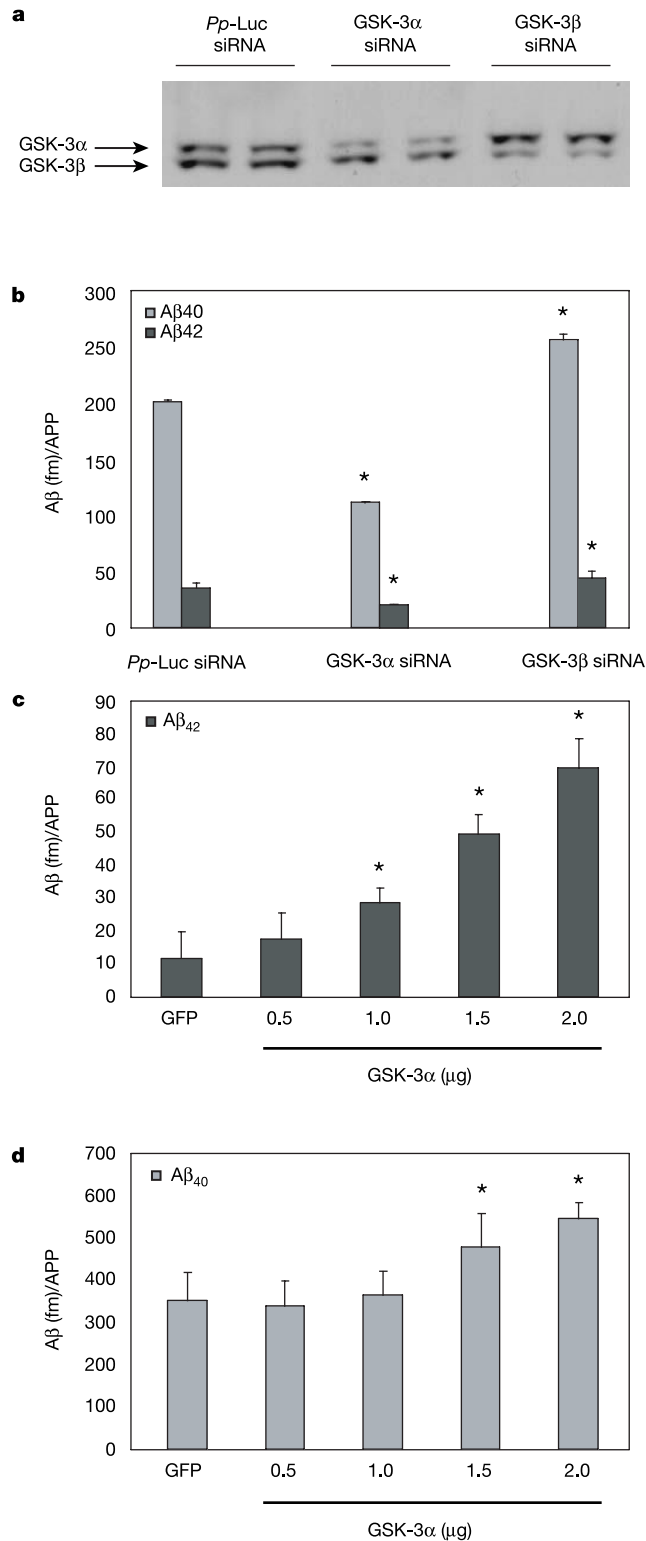


Figure 3 GSK-3 α is required for A β production. **a**, GSK-3 α and GSK-3 β levels were measured by immunoblotting in CHO-APP₆₉₅ cells transfected with siRNAs. siRNA against *Pp-Luc* represents control; GSK-3 α siRNA selectively reduces GSK-3 α protein and GSK-3 β -directed siRNA selectively reduces GSK-3 β . **b**, Secreted A β levels in siRNA transfected cells. The experiment was repeated six times with similar results. **c**, **d**, CHO-APP₆₉₅ cells were transfected with GFP or GSK-3 α , and secreted A β ₄₂ (**c**) and A β ₄₀ (**d**) levels were assessed as in Fig. 1. Error bars represent s.e.m. In panels **b–d**, asterisks indicate a significant difference from control, as determined by one-way ANOVA ($P < 0.05$).

Overexpression of GSK-3 α in CHO-APP₆₉₅ cells increased A β ₄₀ and A β ₄₂ levels in a dose-dependent manner (Fig. 3c, d). Thus, A β production correlated with GSK-3 α expression using both loss-of-function and overexpression approaches. The isoform-specific depletion of GSK-3 by RNAi, together with the overexpression data, shows that GSK-3 α is required for maximal APP processing.

To show that lithium reduces A β production in neurons, primary cultures of mouse embryonic cortical neurons were infected with

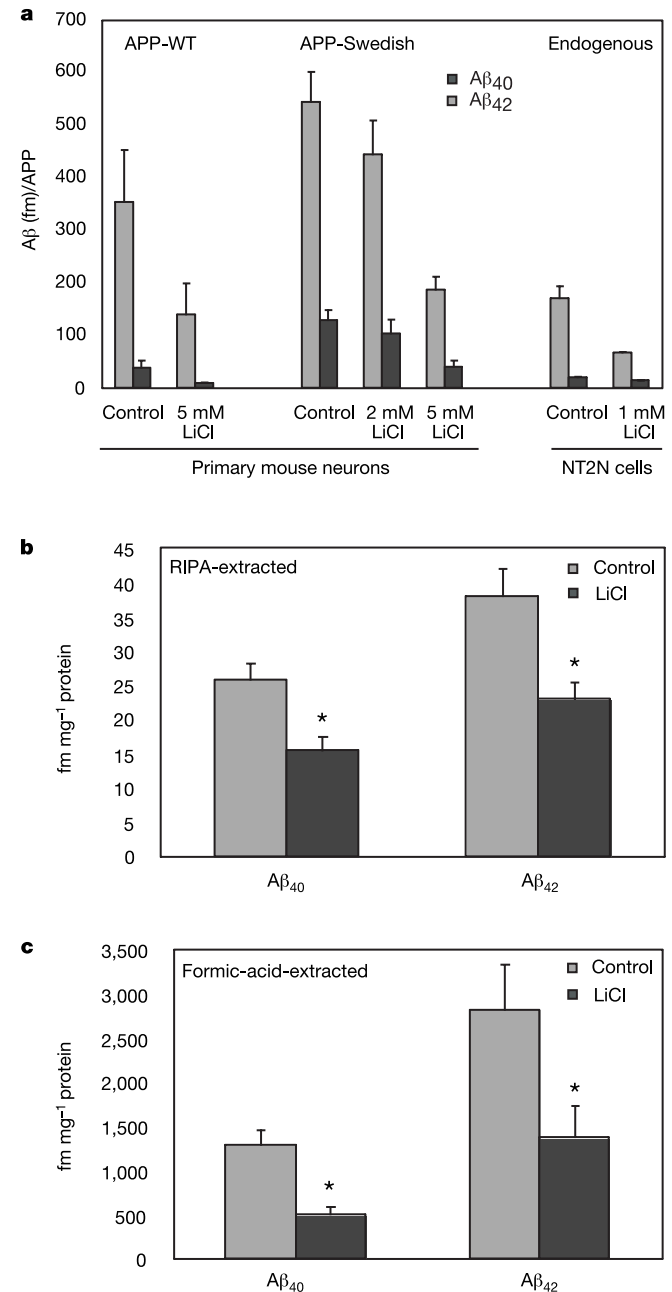


Figure 4 Lithium blocks A β accumulation in cultured neurons and in the brains of mice overproducing A β peptides. **a**, Embryonic cortical neurons were infected with SFV containing wild-type APP (APP-WT) or APP-Swedish (KM670/671NL), then treated with LiCl for 24 h. Secreted A β was measured as described above. **b**, **c**, Three-month-old female mice heterozygous for the APP-Swedish transgene (Tg2576) and the PS1^{P264L} knock-in received 300 mg kg⁻¹ LiCl ($n = 7$) or NaCl ($n = 7$) by gastric gavage once daily for 3 weeks, and cortical A β accumulation was assessed in RIPA-extracted (soluble) and formic-acid-extracted (insoluble) fractions. Serum lithium levels were 0.8–1.2 mM. Error bars represent s.e.m. Asterisks in panel **b** and **c** denote a significant difference from NaCl-treated animals, assessed by one-way ANOVA ($P < 0.05$).

SFV encoding wild-type APP (APP₆₉₅) or the pathogenic APP-Swedish mutation (KM670/671NL), then treated with lithium for 24 h. Production of A β ₄₀ and A β ₄₂ was reduced by 60% and 78%, respectively, for wild-type APP, and to a similar extent for the APP-Swedish mutation (Fig. 4a). A clinically relevant concentration of lithium (1 mM) also reduced the accumulation of endogenous A β by 60% after 3 days of culture in the neuronal NT2N cell line (Fig. 4a).

Lithium also reduces A β production in an animal model of Alzheimer's disease. Mice expressing pathogenic, familial Alzheimer's disease (FAD)-associated forms of APP (APP-Swedish/Tg2576) and PS1 (PS1^{P264L/wt}; ref. 21) were treated with daily lithium gavage for 3 weeks, and the levels of A β peptides were measured in the brain. Under these conditions, serum lithium levels were 0.8–1.2 mM, within the therapeutic range for patients treated with lithium. In the lithium-treated group ($n = 7$), soluble, RIPA-extracted A β ₄₀ and A β ₄₂ levels were each reduced by 40% (Fig. 4b). Furthermore, insoluble A β ₄₀ and A β ₄₂ extractable with formic acid were reduced by 62% and 51%, respectively (Fig. 4c), showing that a therapeutic dose of lithium markedly reduces the tissue level of A β peptides.

In addition to presenilin, nicastrin⁴, APH-1 and PEN-2 (ref. 6) are also required for γ -secretase activity, and loss of any of these components affects both APP and Notch processing. By contrast, lithium does not inhibit Notch processing, indicating that lithium is not a general inhibitor of γ -secretase. Although consensus sites for GSK-3 phosphorylation have been identified in PS1 that are important for PS1 stability, mutation of these sites does not affect γ -secretase activity²². Thus, it is unlikely that GSK-3 α has a role in the biogenesis, stability, or overall activity of the γ -secretase complex. Rather, GSK-3 α might regulate γ -secretase activity toward specific substrates, or access of these substrates to the γ -secretase complex. Certain non-steroidal anti-inflammatory drugs (NSAIDs) also reduce A β ₄₂ levels without affecting Notch cleavage²³. However, the mechanisms of lithium and NSAIDs appear to differ, as NSAIDs shift APP cleavage products from A β ₄₂ to A β ₃₈, whereas lithium reduces overall A β production. This apparent difference in mechanism suggests that combination therapy with lithium and NSAIDs could have an enhanced effect in reducing A β peptide accumulation.

In summary, we have shown that GSK-3 α facilitates APP processing and that lithium inhibits the generation of A β peptides through inhibition of GSK-3 α . In support of this conclusion: (1) lithium reduces A β production in cultured cells and in the brains of mice that overproduce A β peptides; (2) kenpaullone, an alternative GSK-3 α inhibitor, also inhibits A β production; (3) RNAi-mediated depletion of GSK-3 α reduces A β production; and (4) moderate overexpression of GSK-3 α increases A β production. Lithium inhibits the GSK-3-mediated phosphorylation of tau¹², which, in its hyperphosphorylated state, is the main component of neurofibrillary tangles²⁴. Thus, GSK-3 α offers an attractive target for pharmacological agents aimed at reducing the formation of amyloid plaques and neurofibrillary tangles, the pathological hallmarks of Alzheimer's disease. Lithium also protects neurons from proapoptotic stimuli and could therefore reduce neuronal cell death associated with Alzheimer's disease²⁴. Lithium has been used for more than 50 years to treat bipolar disorder, but has a narrow therapeutic window and a higher frequency of side effects in older patients. Thus, although lithium might be considered for the prevention of Alzheimer's disease, especially in younger patients with FAD mutations or Down's syndrome, new agents that specifically target GSK-3 α may prove to be valuable in the treatment of Alzheimer's disease. □

Methods

Plasmids

Human GSK-3 α in pMT2 was provided by J. Woodgett. *Xenopus* β -catenin (obtained

from B. Gumbiner) was subcloned from pCS2 + into pSFV. Notch- Δ E in pCS2MT and NICD in pCS2MT were provided by R. Kopan. Notch- Δ E-V1744K in pCS2MT was created by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene) and confirmed by sequencing. Wild-type APP and APP-Swedish (KM670/671NL) in pSFV were as described previously²⁵. OT-luciferase (OT-Luc) was provided by K. Kinzler and B. Vogelstein.

Cell culture

CHO₆₉₅ cells (obtained from S. Sisodia) were maintained in MEM α + 5% fetal bovine serum (FBS) with added penicillin/streptomycin and glutamine. To generate NT2N neurons, a human embryonic carcinoma cell line (NTera2/D1 or NT2) was grown and maintained as described²⁶. For murine primary embryonic neurons, cortices from embryonic-day-15 mouse brains were isolated and incubated in 0.1% trypsin, HBSS and 0.5 mM EDTA without Ca²⁺ or Mg²⁺ (Invitrogen), and the cells mechanically dissociated using a fire-polished pipette. Cells were plated in DMEM + 10% FBS in poly-D-lysine-coated six-well plates at a density of 10⁶ cells per well. Twenty-four hours after plating, the medium was replaced with DMEM + B27 (Invitrogen) to promote neuronal survival and to inhibit the growth of non-neuronal cells. Neurons were used for experiments after three days in culture.

A β quantification

CHO-APP₆₉₅ cells were plated in six-well dishes at a density of 5 $\times$ 10⁵ cells per well. Stocks of LiCl were prepared in sterile water, whereas DFK-167 (Enzyme Systems Products), kenpaullone (Calbiochem) and roscovitine (Calbiochem) were prepared in dimethylsulphoxide. Drugs were added to cells in fresh medium, and media and cells were collected 24 h later. A β determinations from media were made by sandwich ELISA²⁷. APP levels were quantified by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) followed by immunoblotting with the N-terminal APP antibody Karen²⁸, and visualized using ¹²⁵I-labelled secondary antibodies by PhosphorImager analysis and ImageQuant software (Molecular Dynamics).

Visualization of metabolically labelled APP fragments

CHO₆₉₅ cells plated on six-well dishes were methionine deprived for 30 min by incubation in methionine-free DMEM (Life Technologies) before adding 500 μ Ci ³⁵S-methionine (Perkin Elmer) per ml of DMEM with 5% dialysed FBS (Invitrogen) and 40 mg ml⁻¹ L-proline for 2.5 h. Media was collected and cells (two wells per sample) were rinsed twice with PBS and scraped into RIPA buffer. Lysates were sonicated and centrifuged at 100,000g for 20 min. Media was immunoprecipitated with the antibody BAN-50, and cell lysates were immunoprecipitated with the antibody 2493 using protein A/G agarose beads (Santa Cruz Biotechnologies). Proteins were resolved by electrophoresis on Tris-tricine polyacrylamide gels with a 10/16% step gradient. Gels were fixed in 50% methanol + 5% glycerol, dried and exposed to a PhosphorImager screen.

Viral infections

SFV vectors were used to overexpress β -catenin and APP. SFV was prepared and titred in BHK cells, as previously described²⁹. Cells were infected with SFV in serum-free DMEM at a multiplicity of infection of 10. One hour after infection, the medium was replaced with MEM α + 5% FBS (CHO₆₉₅ cells) or DMEM + B27 supplement (murine primary neurons).

siRNA preparation and transfection

RNA oligonucleotides were synthesized by Dharmacon. Sequences used were: pGL3-luciferase (Pp-Luc) sense, 5'-CUUACGUGAGUACUUCGAdTdT-3'; Pp-Luc antisense, 5'-UCGAAGUACUCAGCGUAAAGdTdT-3'; GSK-3 β sense, 5'-AUCUUUGGAGCCACUGAUdTdT-3'; GSK-3 β antisense, 5'-AAUCAGUGGCUCCAAAGAUdTdT-3'; GSK-3 α sense, 5'-UUCUACUCCAGUGGUGAGdTdT-3'; GSK-3 α antisense, 5'-UCUCACCACUGGAGUAGAdTdT-3'. Double-stranded RNA was produced using conditions described in ref. 9. siRNA was transfected into CHO-APP₆₉₅ cells plated in six-well dishes using GenePORTER (Gene Therapy Systems). siRNA (25 nM) was co-transfected with 2 μ g plasmid DNA and cells were harvested 48 h later. GSK-3 α and GSK-3 β protein levels were examined using an antibody that recognizes both GSK-3 isoforms (Calbiochem).

Notch processing

CHO-APP₆₉₅ cells were plated at a density of 1 $\times$ 10⁵ cells per well in six-well dishes. Cells were transfected with 2 μ g of either Notch-ICV or Notch- Δ E. Twenty-four hours later, the medium was changed and drugs were added. After a further 24 hours, cells were harvested as described above, and immunoblotted with both myc (9E10) antibody or cleaved Notch antibody (Cell Signaling Technologies).

Lithium treatment in mice

Transgenic mice expressing APP-Swedish (Tg2576) were crossed to mice carrying a knock-in of the PS1^{P264L} mutation, as described²⁰, and 3-month-old female Tg2576/PS1^{P264L/WT} mice ($n = 7$) were administered LiCl or NaCl (300 mg kg⁻¹ in 0.4 ml water, by gastric gavage daily for 3 weeks). Animals were allowed free access to food and water, and lithium-treated mice were given free access to a 450-mM NaCl solution. Animals were sacrificed 3 h after the final dose, and A β levels were measured after stepwise extraction of A β from

brain tissues. Briefly, cerebral cortices were lysed in RIPA buffer; after centrifugation, insoluble material was extracted in formic acid, as described elsewhere³⁰. RIPA- and formic-acid-extracted samples were diluted, and A β levels were measured by sandwich ELISA, as above.

Received 21 January; accepted 1 April 2003; doi:10.1038/nature01640.

1. Selkoe, D. J. Presenilin, Notch, and the genesis and treatment of Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **98**, 11039–11041 (2001).
2. Vassar, R. *et al.* β -Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735–741 (1999).
3. De Strooper, B. *et al.* Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature* **391**, 387–390 (1998).
4. Esler, W. P. *et al.* Activity-dependent isolation of the presenilin- γ -secretase complex reveals nicastrin and a γ secretase substrate. *Proc. Natl Acad. Sci. USA* **99**, 2720–2725 (2002).
5. Edbauer, D., Winkler, E., Haass, C. & Steiner, H. Presenilin and nicastrin regulate each other and determine amyloid β -peptide production via complex formation. *Proc. Natl Acad. Sci. USA* **99**, 8666–8671 (2002).
6. Francis, R. *et al.* *aph-1* and *pen-2* are required for Notch pathway signalling, γ -secretase cleavage of β APP, and presenilin protein accumulation. *Dev. Cell* **3**, 85–97 (2002).
7. Soriano, S. *et al.* Presenilin 1 negatively regulates β -catenin/T cell factor/lymphoid enhancer factor-1 signalling independently of β -amyloid precursor protein and notch processing. *J. Cell Biol.* **152**, 785–794 (2001).
8. Yu, G. *et al.* Nicastrin modulates presenilin-mediated *notch/glp-1* signal transduction and β APP processing. *Nature* **407**, 48–54 (2000).
9. Takashima, A. *et al.* Presenilin 1 associates with glycogen synthase kinase-3 β and its substrate tau. *Proc. Natl Acad. Sci. USA* **95**, 9637–9641 (1998).
10. Kang, D. E. *et al.* Presenilin 1 facilitates the constitutive turnover of β -catenin: differential activity of Alzheimer's disease-linked PS1 mutants in the β -catenin-signalling pathway. *J. Neurosci.* **19**, 4229–4237 (1999).
11. Kang, D. E. *et al.* Presenilin couples the paired phosphorylation of β -catenin independent of axin: implications for β -catenin activation in tumorigenesis. *Cell* **110**, 751–762 (2002).
12. Phiel, C. J. & Klein, P. S. Molecular targets of lithium action. *Annu. Rev. Pharmacol. Toxicol.* **41**, 789–813 (2001).
13. Klein, P. S. & Melton, D. A. A molecular mechanism for the effect of lithium on development. *Proc. Natl Acad. Sci. USA* **93**, 8455–8459 (1996).
14. Sun, X. *et al.* Lithium inhibits amyloid secretion in COS7 cells transfected with amyloid precursor protein C100. *Neurosci. Lett.* **321**, 61–64 (2002).
15. De Strooper, B. *et al.* A presenilin-1-dependent γ -secretase-like protease mediates release of Notch intracellular domain. *Nature* **398**, 518–522 (1999).
16. Wolfe, M. S. *et al.* Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and γ -secretase activity. *Nature* **398**, 513–517 (1999).
17. Schroeter, E. H., Kisslinger, J. A. & Kopan, R. Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature* **393**, 382–386 (1998).
18. Bain, J., McLaughlan, H., Elliott, M. & Cohen, P. The specificities of protein kinase inhibitors—an update. *Biochem. J.* **371**, 199–204 (2003).
19. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
20. Hoeflich, K. P. *et al.* Requirement for glycogen synthase kinase-3 β in cell survival and NF- κ B activation. *Nature* **406**, 86–90 (2000).
21. Siman, R. *et al.* Presenilin-1 P264L knock-in mutation: differential effects on A β production, amyloid deposition, and neuronal vulnerability. *J. Neurosci.* **20**, 8717–8726 (2000).
22. Kirschenbaum, F., Hsu, S. C., Cordell, B. & McCarthy, J. V. Substitution of a glycogen synthase kinase-3 β phosphorylation site in presenilin 1 separates presenilin function from β -catenin signaling. *J. Biol. Chem.* **276**, 7366–7375 (2001).
23. Weggen, S. *et al.* A subset of NSAIDs lower amyloidogenic A β 42 independently of cyclooxygenase activity. *Nature* **414**, 212–216 (2001).
24. Alvarez, G. *et al.* Regulation of tau phosphorylation and protection against β -amyloid-induced neurodegeneration by lithium. Possible implications for Alzheimer's disease. *Bipolar Disord.* **4**, 153–165 (2002).
25. Forman, M. S., Cook, D. G., Leight, S., Doms, R. W. & Lee, V. M.-Y. Differential effects of the Swedish mutant amyloid precursor protein on β -amyloid accumulation and secretion in neurons and nonneuronal cells. *J. Biol. Chem.* **272**, 32247–32253 (1997).
26. Pleasure, S. J., Page, C. & Lee, V. M.-Y. Pure, postmitotic, polarized human neurons derived from NTera 2 cells provide a system for expressing exogenous proteins in terminally differentiated neurons. *J. Neurosci.* **12**, 1802–1815 (1992).
27. Suzuki, N. *et al.* An increased percentage of long amyloid β protein secreted by familial amyloid β protein precursor (BAPP717) mutants. *Science* **264**, 1336–1340 (1994).
28. Turner, R. S., Suzuki, N., Chyung, A. S., Younkin, S. G. & Lee, V. M.-Y. Amyloids β 40 and β 42 are generated intracellularly in cultured human neurons and their secretion increases with maturation. *J. Biol. Chem.* **271**, 8966–8970 (1996).
29. Cook, D. G. *et al.* Alzheimer's A β (1–42) is generated in the endoplasmic reticulum/intermediate compartment of NT2N cells. *Nature Med.* **3**, 1021–1023 (1997).
30. Wilson, C. A., Doms, R. W., Zheng, H. & Lee, V. M.-Y. Presenilins are not required for A β 42 production in the early secretory pathway. *Nature Neurosci.* **5**, 849–855 (2002).

Acknowledgements We wish to thank R. Doms for critical reading of the manuscript, and T. Kadesch, J. Yang, G. Wertheim, K. Phiel and members of the Klein, Lee and Doms laboratories for helpful discussions. We are grateful to J. Woodgett, R. Kopan, B. Gumbiner, K. Kinzler and B. Vogelstein for plasmids, S. Sisodia for CHO-APP₆₉₅ cells, and D. Flood for PS1^{P264L} knock-in mice. Monoclonal antibodies for the A β sandwich ELISA were provided by N. Suzuki and Tekeda Pharmaceuticals. C.A.W. was a Howard Hughes Predoctoral Fellow. This work was supported by grants from the National Institute on Aging (to V.M.-Y.L.) and the National Institute of Mental Health (to P.S.K.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.S.K. (pklein@mail.med.upenn.edu).

Caenorhabditis elegans early embryogenesis and vulval morphogenesis require chondroitin biosynthesis

Ho-Yon Hwang^{*}, Sara K. Olson[†], Jeffrey D. Esko[†] & H. Robert Horvitz^{*}

^{*} Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Room 68-425, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

[†] Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, 9500 Gilman Drive, CMM-East 1055, La Jolla, California 92093, USA

Defects in glycosaminoglycan biosynthesis disrupt animal development and can cause human disease^{1–4}. So far much of the focus on glycosaminoglycans has been on heparan sulphate. Mutations in eight *squashed vulva* (*sqv*) genes in *Caenorhabditis elegans* cause defects in cytokinesis during embryogenesis and in vulval morphogenesis during postembryonic development^{5,6}. Seven of the eight *sqv* genes have been shown to control the biosynthesis of the glycosaminoglycans chondroitin and heparan sulphate^{6–11}. Here we present the molecular identification and characterization of the eighth gene, *sqv-5*. This gene encodes a bifunctional glycosyltransferase that is probably localized to the Golgi apparatus and is responsible for the biosynthesis of chondroitin but not heparan sulphate. Our findings show that chondroitin is crucial for both cytokinesis and morphogenesis during *C. elegans* development.

Glycosaminoglycans (GAGs) or mucopolysaccharides have been of great interest to biologists for decades^{2–4,12–15}. Analyses of mutations that cause developmental defects in *Drosophila melanogaster* have shown that GAG biosynthesis, in particular the synthesis of heparan sulphate (HS), is important for intercellular signalling mediated by the Wingless, Hedgehog and fibroblast growth factor pathways². In addition, mutations in GAG biosynthesis have been implicated in human diseases, including a progeroid variant of the connective tissue disorder Ehlers–Danlos syndrome¹ and hereditary multiple exostoses³, which is characterized by inappropriate chondrocyte proliferation and bone growth. Much of the focus on GAGs has been on HS, in part because of the many ligands that it binds, its action in growth factor signalling, and its role in *Drosophila* development^{2,14}. Studies of chondroitin sulphate (CS), another chief class of GAGs in vertebrates, have focused on the development of cartilage, tendon and bone⁴.

The *C. elegans* genes *sqv-1* to *sqv-8* are important for both embryonic development and postembryonic vulval morphogenesis⁵. The progeny of mutants homozygous for strong loss-of-

function *sqv* mutations die during embryogenesis, with most arresting at the one-cell stage. This arrest is caused by a defect in the initiation of cytokinesis. This defect may be caused by a failure to form a fluid-filled extracellular space between the plasma membrane and the eggshell⁶. During the L4 larval stage, *sqv-1* to *sqv-8* mutants fail to expand the extracellular space of the vulva, which is the opening through which sperm and eggs pass in adult hermaphrodites. These mutants form a partially functional vulva but are normal in vulval cell proliferation, migration and fusion. We previously proposed that during both embryogenesis and vulval morphogenesis in *C. elegans*, GAGs that are attached to extracellular matrices drive the formation of fluid-filled extracellular spaces⁶. In sea urchin fertilization, the secretion of GAGs is thought to cause the swelling of a fluid-filled space between the vitelline envelope and the plasma membrane¹⁶.

The molecular identities of seven of the *sqv* genes indicate a defect in the biosynthesis of two types of GAG present in *C. elegans*, HS and chondroitin^{6–11}. SQV-4 (UDP-glucose dehydrogenase) synthesizes UDP-glucuronic acid (UDP-GlcA) in the cytoplasm¹⁰, which is translocated into the lumen of the Golgi apparatus by the SQV-7 nucleotide-sugar transporter⁹. SQV-7 also translocates UDP-galactose and UDP-N-acetylgalactosamine⁹. SQV-1 catalyses the decarboxylation of UDP-GlcA⁶, forming the first nucleotide-sugar donor required for GAG biosynthesis, UDP-xylose. In the lumen of the Golgi apparatus, UDP-xylose, UDP-galactose and UDP-GlcA are used as substrates of the SQV-6 xylosyltransferase¹¹, the SQV-3 galactosyltransferase I (refs 7,8), the SQV-2 galactosyltransferase II (ref. 11) and the SQV-8 glucuronosyltransferase I (refs 7,8) to build onto the protein core the GAG linkage tetrasaccharide (GlcA β 1,3Gal β 1,3Gal β 1,4Xyl β -O-serine) on which GAG backbones polymerize. Evidence that four such glycosylation reactions

are essential for mammalian GAG polymerization has been obtained from studies of mutant hamster cell lines¹⁴. Mutation in the human homologue of the *sqv-3* galactosyltransferase I gene has been implicated as the cause of a progeroid variant of the connective-tissue disorder Ehlers–Danlos syndrome^{17,18}, which is characterized by loose skin and hypermobile joints. It seems likely that homologues of other *sqv* genes are involved in similar disorders.

By physically mapping chromosomal deletions, we localized *sqv-5* to a region of roughly 200 kilobases (kb) between *fog-3* and the left endpoint of *qDf10* (Methods and Supplementary Fig. S1). A *Bam*HI–*Pst*I fragment of 18,446 nucleotides of cosmid K09A8 containing a single complete predicted gene (*T24D1.1*) rescued the *sqv-5* mutant phenotype. Introducing a nonsense or frameshift mutation in *T24D1.1* eliminated this rescuing activity. The *sqv-5* *n3039* allele is a nonsense mutation in the *T24D1.1* open reading frame (ORF). We isolated a mutation (*n3611*) that deleted most of the *T24D1.1* ORF and caused the same *sqv* mutant phenotype as that of *sqv-5*(*n3039*) worms. We conclude that *sqv-5* corresponds to *T24D1.1*.

We found three discrepancies between our DNA sequencing results and those of the *C. elegans* Sequencing Consortium, one of which caused us to modify the predicted gene structure of *T24D1.1* to that shown in Supplementary Fig. S1. From the sequences of *sqv-5* complementary DNA and products from the 5' rapid amplification of cDNA ends (RACE) products, we identified two alternatively spliced forms of *sqv-5* cDNAs, which encode predicted proteins of 734 and 736 amino acids. We detected a transcript of 3.6 kb on a northern blot (Supplementary Fig. S2), consistent with the size predicted by our cDNA and 5' RACE results.

Of the 734 amino acids in the short form of the SQV-5 protein, 270 (37%) are identical to a human CS synthase (ref. 19, Fig. 1 and Supplementary Table S1). We also identified and determined the

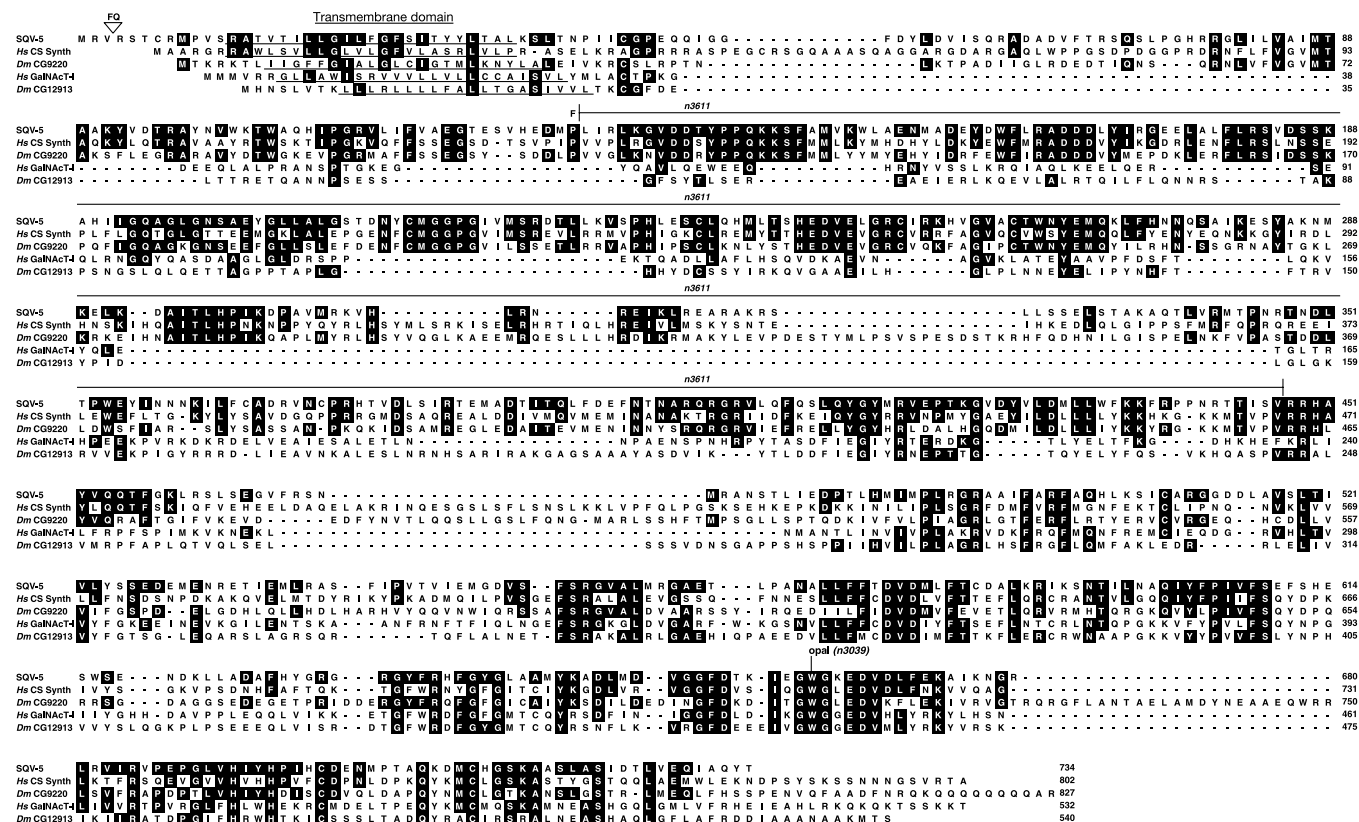


Figure 1 Sequence alignment of SQV-5 and homologues. Human CS synthase (*Hs* CS Synth), *Drosophila melanogaster* CS synthase homologue (*Dm* CG9220), human CS GalNAcT-I (*Hs* GalNAcT-I) and *Drosophila* CS GalNAcT-I homologue (*Dm* CG12913) are compared with SQV-5. The extent of the *sqv-5*(*n3611*) deletion and the location of the

sqv-5(*n3039*) nonsense allele are indicated. Two of the six cloned 5' RACE products contained six extra nucleotides at the 5' end of the second exon, reflecting an alternatively spliced mRNA; the addition of two amino acids (FQ) in the longer alternatively spliced form of SQV-5 is indicated.

sequences of a cDNA from a *Drosophila melanogaster* gene that is predicted to encode a protein of 832 amino acids, of which 270 are identical to SQV-5 (37%). SQV-5 is less similar to the human CS *N*-acetylglucosaminyltransferase I (GalNAcT-I)²⁰, with which it shares 109 amino acids (20% identity). However, because SQV-5 is the only protein in the *C. elegans* genome with extensive similarity to the human CS GalNAcT-I, we considered SQV-5 also to be a candidate CS GalNAcT-I. By contrast, the *Drosophila* genome contains a second gene that probably encodes an orthologue of the human CS GalNAcT-I (37% identity) (Fig. 1 and Supplementary Table S1). All five proteins contain a single predicted transmembrane domain near the amino terminus, consistent with a type II transmembrane topology. Such N-terminal transmembrane domains are typical of glycosyltransferases, which have active sites that face the lumen of the Golgi apparatus.

In vertebrates, CS synthase catalyses the alternating, stepwise addition of GlcA and *N*-acetylglucosamine (GalNAc) to the nascent chain, resulting in the polymerization of the CS backbone. Protein extracts prepared from whole worms that were homozygous for the *sqv-5(n3611)* null allele, heterozygous for *sqv-5(n3611)* or wild-type were assayed for glucuronosyltransferase (GlcAT-II) and *N*-acetylglucosaminyltransferase (GalNAcT-II) activities of CS synthase. Significant enzymatic activities were observed in the extracts of the wild-type worms in both assays using chemically desulphated CS as the acceptor (Fig. 2a, b). Worms heterozygous for *sqv-5(n3611)* contained slightly more than a half of the enzymatic activity observed in wild-type worms, and no significant enzymatic activity above the negative control was observed in the extracts of the homozygous mutant worms.

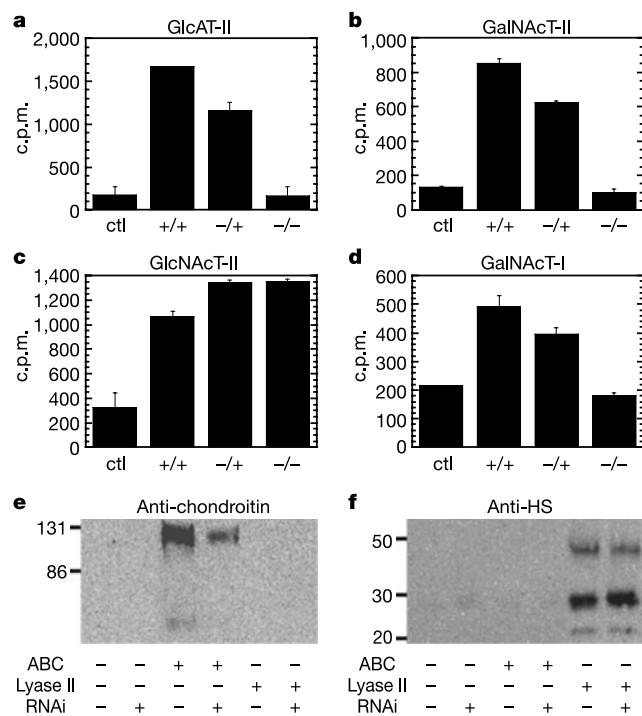


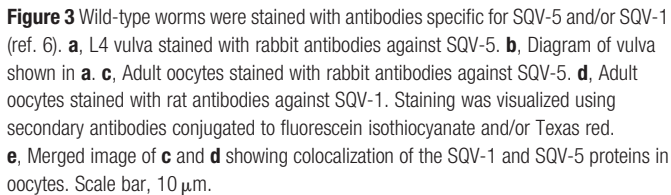
Figure 2 Glycosyltransferase activities were assayed using cell-free extracts prepared from wild type (+/+) and *sqv-5(n3611)* heterozygotes (-/+) and homozygotes (-/-). Activities measured without an acceptor (ctl) are also shown. Results are the mean $\pm$ s.e.m. from representative experiments. **a**, GlcAT-II assay. **b**, GalNAcT-II assay. **c**, GlcNAcT-II assay. **d**, GalNAcT-I assay. **e**, **f**, Intact proteoglycans were purified from vector (-) or *sqv-5* (+) RNAi-treated worms and treated with chondroitinase ABC and/or heparin lyase II or buffer. Western blot analysis was done with monoclonal antibodies specific for chondroitin (**e**) or HS (**f**), which were visualized using secondary antibodies conjugated to horseradish peroxidase.

About half the enzymatic activity was observed in assays consisting of half wild-type and half *sqv-5* null extracts (wild type, $1,466 \pm 144$; *sqv-5*, 105 ± 35 ; wild type + *sqv-5*, 809 ± 57 , control, 68 ± 24 c.p.m.; mean $\pm$ s.d.), indicating that the mutant worms do not contain an inhibitor of the enzyme. These findings establish that *sqv-5* controls chondroitin synthase activity in *C. elegans* and, in combination with our sequence data, show that the SQV-5 protein has synthase activity. By contrast, we observed no difference in *N*-acetylglucosaminyltransferase II (GlcNAcT-II) activity, which is required for HS synthesis, in extracts of *sqv-5(n3611)* mutants (Fig. 2c), suggesting that chondroitin but not HS biosynthesis is disrupted in *sqv-5* mutants.

The addition of the first GalNAc residue to the linkage tetrasaccharide is thought to be catalysed by a different enzyme to that involved in CS backbone polymerization^{19,20}. This reaction can be assayed using glucuronic acid β 1,3galactose-O-naphthalenemethanol as the acceptor instead of desulphated CS. Using this assay, we observed significant enzymatic activity in extracts of wild-type worms (Fig. 2d). About one-half of the enzymatic activity above that of the negative control was present in extracts of worms heterozygous for *sqv-5(n3611)*, whereas homozygous null worms lacked activity. These findings indicate that SQV-5 functions in both the initiation and the elongation of chondroitin chains. It is notable that SQV-5, which has a similar degree of amino acid sequence identity to the human GalNAcT-I as to the human CS synthase (Supplementary Table S1), has this additional GalNAcT-I activity, as the human CS synthase apparently lacks this activity¹⁹.

To assay changes in chondroitin and HS content in *sqv-5* mutants, we isolated GAGs from wild-type and *sqv-5(n3611)* adult hermaphrodites. The amount of chondroitin was reduced from 182 ± 52 fmol per worm (mean $\pm$ s.d.) in wild type to 24 ± 20 fmol per worm in *sqv-5* mutants. The amount of HS was below the limits of detection; the relative amount of HS is 150- to 250-fold less than chondroitin in *C. elegans*^{21,22}. We also suppressed *sqv-5* function by RNA-mediated interference (RNAi) achieved by feeding the worms dsRNA²³. *sqv-5* RNAi-treated L4 larvae had reduced vulval extracellular spaces reminiscent of mutants homozygous for a weak loss-of-function mutation in other *sqv* genes (Supplementary Fig. S3). The *sqv-5* RNAi-treated adults treated were reduced in brood size (29 ± 23 (mean $\pm$ s.d.), $n = 47$) as compared with vector RNA-treated adults (264 ± 33 , $n = 31$). These observations suggest that *sqv-5* function is incompletely suppressed in these worms, because worms homozygous for either *sqv-5* mutant allele have average brood sizes of zero. Intact proteoglycans were isolated from L4 larvae that had been treated with *sqv-5* or vector RNAi and digested with either chondroitinase ABC or heparin lyase II. These enzyme treatments of chondroitin and HS resulted in 'stub' oligosaccharides, consisting of the linkage tetrasaccharide and one disaccharide repeat containing a terminal 4,5-unsaturated uronic acid, which were recognized by monoclonal antibodies against either chondroitin or HS, respectively. Western blot analysis using the anti-chondroitin antibody detected a major proteoglycan band with a relative molecular mass of about 120,000 ($M_r \approx 120K$) and several minor bands. The main band of 120K was reduced 5–8-fold in *sqv-5* RNAi-treated worms as compared with vector RNA-treated worms (Fig. 2e, f). Western blot analysis using the anti-HS antibody detected three proteoglycans of 24K, 30K and 48K, which did not vary in amount between vector and *sqv-5* RNAi-treated worms. Thus, loss of *sqv-5* function selectively reduces chondroitin quantities.

To study the expression and subcellular localization of the SQV-5 protein, we generated affinity-purified rabbit polyclonal antibodies against a SQV-5 and glutathione S-transferase (GST) fusion protein. Anti-SQV-5 antibodies stained several punctate foci in the cytoplasm of the vulva, the uterus and oocytes (Fig. 3a–c). This punctate staining was not seen in worms homozygous for the *sqv-5(n3611)* null allele (data not shown).



Our molecular identification of *sqv-5* defines the last step in the *C. elegans* biosynthetic pathway for chondroitin and suggests that defects in the biosynthesis of chondroitin account for the embryonic and vulval defects caused by mutations in all *sqv* genes (Fig. 4). By contrast,

In vertebrates, large amounts of CS are secreted into the extracellular matrix, where CS has a structural role and binds to ligands such as type I collagen¹³. The ability of chondroitin to interact with water, to cause swelling and to generate osmotic pressure on its surroundings could be responsible for its biological effects¹², including the expansion of the extracellular spaces of the *C. elegans* embryo and vulva. Studies of sea urchin gastrulation have led to the proposal that the secretion of CS proteoglycan can result in the hydration of the extracellular matrix and cause epithelial invagination²⁴. Our studies provide support for such a mechanism. We cannot exclude other possibilities, however, such as mechanisms involving adhesion, cytoskeletal rearrangement or intercellular signalling^{5,6}. Whatever the mechanism of chondroitin action, our findings show the importance of chondroitin in cytokinesis, early embryogenesis and epithelial morphogenesis in *C. elegans* and support the hypothesis that CS, like HS, has a broad and important role in development and disease. □

sqv-5 mapping

We obtained Unc non-Vul and Vul non-Unc progeny from *unc-29(e1072) lin-11(n566)/sqv-5(n3039)* hermaphrodites. Five of ten Unc non-Vul progeny carried *sqv-5(n3039)*, and three of ten Vul non-Unc progeny carried *sqv-5(n3039)*, indicating that *sqv-5(n3039)* was located between *unc-29* and *lin-11* and to the left of *lin-11*. We examined the vulval phenotype of worms of the genotype *ces-1(n703) Df[sqv-5(n3039)]*, using *qDf5*, *qDf7*, *qDf8*, *qDf9* and *qDf10*. All worms were Sqv, except for *ces-1(n703) qDf5/sqv-5(n3039)*. Because *qDf5*, *qDf7*, *qDf8*, *qDf9* and *qDf10* delete *fog-3*, but only *qDf5* and *qDf7* delete *lin-11* (ref. 25), *sqv-5* maps to the left of *fog-3*. Using single *qDf10* eggs, we amplified genomic DNA sequence corresponding to the cosmids K10C3 and C03C11. A polymerase chain reaction product of the expected length was amplified for K10C3 but not for C03C11 ($n = 10$), thus placing the left end point of *qDf10* between K10C3 and C03C11 (Supplementary Fig. S1).

***sqv-5* cDNA**

We determined the sequences of two cDNA clones, yk20d7 and yk21g9, corresponding to *T24D1.1*, and of six 5' RACE products derived from mixed-stage RNA. The 5' RACE products contained a 5' SL1 *trans*-spliced leader, which is found at the 5' end of many *C. elegans* transcripts. The *sqv-5* cDNA contained a 417-bp 5' UTR, a 2,202-bp ORF and a 657-bp 3' UTR sequence. We identified two alternatively spliced forms of the transcript by 5' RACE. Two of six cloned 5' RACE products represented a longer spliced form containing six additional base pairs at the 5' end of the second exon (Supplementary Fig. S4).

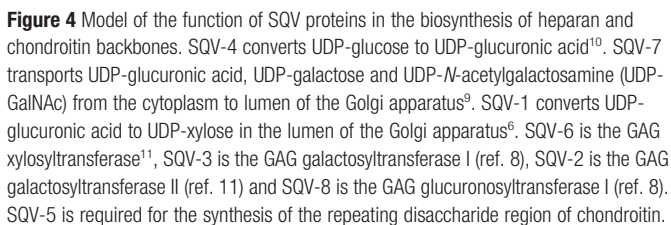
Deletion allele of *sqv-5*

We isolated the deletion mutation *sqv-5(n3611)* from a library of worms mutagenized with ultraviolet illumination and trimethylpsoralen¹⁶. Mutant worms containing *sqv-5(n3611)* were backcrossed to the wild-type strain N2 six times. The deletion of 1,641 bp in *sqv-5(n3611)* removes nucleotides 1,661–3,301 of the *sqv-5* genomic DNA sequence file (GenBank accession code AY241925). *sqv-5(n3611)* is predicted to encode a truncated SQV-5 lacking 385 amino acids (amino acids 130–447) in the middle of SQV-5 and an alanine to phenylalanine substitution at amino acid 129.

Glycosyltransferase assays

We picked *sqv-5(n3611)*, *sqv-5(n3611)/hT2* and wild-type N2 hermaphrodites as L4 larvae by visual examination of the vulva using a dissecting microscope. The worms were allowed to grow for 23–27 h at 22 °C, and were then frozen in 50 mM Tris, pH 7.5 and stored at –70 °C. Samples were sonicated in 0.05% Triton X-100, 50 mM Tris and centrifuged at 15,000g for 10 min. The protein content of the cleared supernatant was assessed by the Bradford assay and portions of the extracts were used for the following assays. The chondroitin acceptor was prepared by desulphation of shark cartilage chondroitin-4-sulphate²⁷. N-acetylheparosan was prepared from *Escherichia coli* K5 (ref. 28) and the disaccharide GlcAβ1,3Galβ-O-naphthalenemethanol was synthesized²⁸.

The GlcAT-II activity of chondroitin synthase was measured by mixing 1.3×10^5 c.p.m. of UDP-[1- ^3H]glucuronic acid donor (20 Ci mmol^{-1}), $6 \mu\text{g}$ of β -glucuronidase-treated chondroitin acceptor and $3 \mu\text{g}$ of worm extract in a volume of $25 \mu\text{l}$ containing 0.05% Triton X-100, 10 mM MnCl_2 , $100 \mu\text{M}$ ATP and 25 mM MES buffer, pH 6.5. The GalNAcAT-II activity of chondroitin synthase was detected by mixing 3×10^5 c.p.m. of UDP-[1- ^3H]GalNAc donor ($38.5 \text{ Ci mmol}^{-1}$), $12 \mu\text{g}$ of chondroitin acceptor and $15 \mu\text{g}$ of



worm extract in 25 μ l containing 0.05% Triton X-100, 10 mM MnCl₂ and 25 mM MES buffer, pH 6.5. The GlcNAcT-II activity of HS polymerase was assayed by mixing 5 μ Ci of UDP-[6-³H]GlcNAc, 12 μ g of N-acetylheparosan acceptor and 10 μ g of worm extract in 25 μ l containing 20 mM MnCl₂, 0.45% Triton X-100 and 25 mM MOPS buffer, pH 6.5. We incubated GlcAT-II, GlcNAcT-II and GalNAcT-II reactions for 2 h at 25 °C and separated the products from free nucleotide-sugars by DEAE-Sephacel (Pharmacia)²⁹.

GalNAcT-I activity was assayed as described for GalNAcT-II, except that 5×10^5 c.p.m. of UDP-[1-³H]GalNAc donor and 12 mM GlcA β 1,3Gal β -O-naphthalenemethanol acceptor were used. Reactions were incubated 3 h at 25 °C, and products were separated from free nucleotide-sugars using a Sep-Pak C18 cartridge (Waters Corp.) as described²⁸. Reactions were linear with time and amount of protein.

Chondroitin and HS characterization

Between 220 and 250 *sqv-5(n3611)* and wild-type worms were collected as described for glycosyltransferase assays, except that the worms were lyophilized and homogenized in acetone. Free GAGs were isolated by alkali extraction as described³⁰, except that they were extracted overnight in 0.5 M NaOH, 1 M NaBH₄ at 4 °C, and neutralized with 1 M HCl. Chondroitin was digested with 20 mU of chondroitinase ABC (Seikagaku) and analysed by high-performance liquid chromatography with post-column derivatization of the disaccharides²².

RNAi was done essentially as described²³, except that L1-stage hermaphrodites were placed onto Petri plates containing bacteria and grown for 44–48 h at 20 °C before being collected as L4-stage hermaphrodites. We isolated intact proteoglycans by anion-exchange chromatography as described³⁰, except that we sonicated the worms in 0.5% Triton X-100 and protease inhibitor cocktail (Sigma). For western blots, intact proteoglycans were digested with chondroitinase ABC or heparin lyase II and incubated with monoclonal antibodies to chondroitin (1-B-5, Seikagaku) or HS (F69-3G10, Seikagaku).

Anti-SQV-5 antibodies

The *sqv-5* ORF was cloned into vectors pGEX-4T3 and pMAL-c2 to generate GST–SQV-5 and maltose-binding protein (MBP)–SQV-5 fusion proteins, respectively. The GST–SQV-5 and MBP–SQV-5 fusion proteins were purified by isolating insoluble proteins from inclusion bodies, followed by SDS–PAGE and electro-elution. We injected GST–SQV-5 into two rabbits (Covance). Anti-SQV-5 antibodies were affinity-purified by binding and elution from MBP–SQV-5 fusion protein, as described¹⁰.

Received 28 January; accepted 14 March 2003; doi:10.1038/nature01634.

1. Quentin, E., Gladen, A., Roden, L. & Kresse, H. A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. *Proc. Natl Acad. Sci. USA* **87**, 1342–1346 (1990).
2. Perrimon, N. & Bernfield, M. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* **404**, 725–728 (2000).
3. Zak, B. M., Crawford, B. E. & Esko, J. D. Hereditary multiple exostoses and heparan sulfate polymerization. *Biochim. Biophys. Acta* **1573**, 346–355 (2002).
4. Schwartz, N. B. & Domowicz, M. Chondrodysplasias due to proteoglycan defects. *Glycobiology* **12**, 57R–68R (2002).
5. Herman, T., Hartwig, E. & Horvitz, H. R. *sqv* mutants of *Caenorhabditis elegans* are defective in vulval epithelial invagination. *Proc. Natl Acad. Sci. USA* **96**, 968–973 (1999).
6. Hwang, H.-Y. & Horvitz, H. R. The SQV-1 UDP-glucuronic acid decarboxylase and the SQV-7 nucleotide-sugar transporter may act in the Golgi apparatus to affect *C. elegans* vulval morphogenesis and embryonic development. *Proc. Natl Acad. Sci. USA* **99**, 14218–14223 (2002).
7. Herman, T. & Horvitz, H. R. Three proteins involved in *Caenorhabditis elegans* vulval invagination are similar to components of a glycosylation pathway. *Proc. Natl Acad. Sci. USA* **96**, 974–979 (1999).
8. Bulik, D. A. *et al.* *sqv-3*, *-7*, and *-8*, a set of genes affecting morphogenesis in *Caenorhabditis elegans*, encode enzymes required for glycosaminoglycan biosynthesis. *Proc. Natl Acad. Sci. USA* **97**, 10838–108343 (2000).
9. Berninsone, P., Hwang, H. Y., Zemtsva, I., Horvitz, H. R. & Hirschberg, C. B. SQV-7, a protein involved in *Caenorhabditis elegans* epithelial invagination and early embryogenesis, transports UDP-glucuronic acid, UDP-N-acetylgalactosamine, and UDP-galactose. *Proc. Natl Acad. Sci. USA* **98**, 3738–3743 (2001).
10. Hwang, H.-Y. & Horvitz, H. R. The *C. elegans* vulval morphogenesis gene *sqv-4* encodes a UDP-glucose dehydrogenase that is temporally and spatially regulated. *Proc. Natl Acad. Sci. USA* **99**, 14224–14229 (2002).
11. Hwang, H.-Y., Olson, S. K., Brown, J. R., Esko, J. D. & Horvitz, H. R. The *C. elegans* genes *sqv-2* and *sqv-6*, which are involved in vulval morphogenesis, encode glycosaminoglycan galactosyltransferase II and xylosyltransferase. *J. Biol. Chem.* **278**, 11735–11738 (2003).
12. Comper, W. D. & Laurent, T. C. Physiological function of connective tissue polysaccharides. *Physiol. Rev.* **58**, 255–315 (1978).
13. Ruoslahti, E. Structure and biology of proteoglycans. *Annu. Rev. Cell Biol.* **4**, 229–255 (1988).
14. Esko, J. D. & Selleck, S. B. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* **71**, 435–471 (2002).
15. Silbert, J. E. & Sugumar, G. Biosynthesis of chondroitin/dermatan sulfate. *Intl Union Biochem. Mol. Biol. Life* **54**, 177–186 (2002).
16. Austin, C. R. *Fertilization* (New Jersey, Prentice Hall, Englewood Cliffs, 1965).
17. Almeida, R. *et al.* Cloning and expression of a proteoglycan UDP-galactose: β -xylose β 1,4-galactosyltransferase I. *J. Biol. Chem.* **274**, 26165–26171 (1999).
18. Okajima, T., Fukumoto, S., Furukawa, K. & Urano, T. Molecular basis for the progeroid variant of Ehlers–Danlos syndrome. *J. Biol. Chem.* **274**, 28841–28844 (1999).
19. Kitagawa, H., Uyama, T. & Sugahara, K. Molecular cloning and expression of a human chondroitin synthase. *J. Biol. Chem.* **276**, 38721–38726 (2001).
20. Uyama, T., Kitagawa, H., Tamura, J. I., & Sugahara, K. Molecular cloning and expression of human chondroitin N-acetylglucosaminyltransferase. *J. Biol. Chem.* **277**, 8841–8846 (2002).
21. Yamada, S. *et al.* Demonstration of glycosaminoglycans in *Caenorhabditis elegans*. *FEBS Lett.* **459**, 327–331 (1999).

22. Toyoda, H., Kinoshita-Toyoda, A. & Selleck, S. B. Structural analysis of glycosaminoglycans in *Drosophila* and *Caenorhabditis elegans* and demonstration that *tout-velu*, a *Drosophila* gene related to EXT tumour suppressors, affects heparan sulfate *in vivo*. *J. Biol. Chem.* **275**, 2269–2275 (2000).
23. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).
24. Lane, M. C., Koehl, M. A., Wilt, F. & Keller, R. A role for regulated secretion of apical extracellular matrix during epithelial invagination in the sea urchin. *Development* **117**, 1049–1060 (1993).
25. Ellis, R. E. & Kimble, J. The *fog-3* gene and regulation of cell fate in the germ line of *Caenorhabditis elegans*. *Genetics* **139**, 561–577 (1995).
26. Jansen, G., Hazendonk, E., Thijssen, K. L. & Plasterk, R. H. Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nature Genet.* **17**, 119–121 (1997).
27. Nagasawa, K., Inoue, Y. & Kamata, T. Solvolytic desulfation of glycosaminoglycansulfates with dimethyl sulfoxide containing water or methanol. *Carbohydr. Res.* **58**, 47–55 (1977).
28. Fritz, T. A., Gabb, M. M., Wei, G. & Esko, J. D. Two N-acetylglucosaminyltransferases catalyze the biosynthesis of heparan sulfate. *J. Biol. Chem.* **269**, 28809–28814 (1994).
29. Wei, G. *et al.* Location of the glucuronosyltransferase domain in the heparan sulfate copolymerase EXT1 by analysis of Chinese hamster ovary cell mutants. *J. Biol. Chem.* **275**, 27733–27740 (2000).
30. Esko, J. D. *in Current Protocols in Molecular Biology* (ed. Ausubel, F.) 17.2.1–17.2.9 (John Wiley and Sons, New York, 1993).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank B. Castor for help with DNA sequencing; Y. Kohara for the cDNA clones yk20dy and yk21g9; A. Coulson for cosmid; J. Brown for the GlcA β 1,3Gal-O-NM acceptor; B. Zak for N-acetylheparosan acceptor; the Glycotechnology Core at UCSD (supported by an NIH grant) for the preparation of UDP-[1-³H]GlcA; and B. Galvin and I. Perez de la Cruz for reading the manuscript. This work was supported by NIH grants (to H.R.H. and J.D.E.). S.K.O. was supported by an NIH training grant. H.R.H. is an Investigator of the Howard Hughes Medical Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.R.H. (horvitz@mit.edu). The *sqv-5* and *D. melanogaster* chondroitin synthase sequences have been deposited in GenBank under accession codes AY241925 and AY24196, respectively.

Chondroitin proteoglycans are involved in cell division of *Caenorhabditis elegans*

Souhei Mizuguchi*†‡, Toru Uyama‡, Hiroshi Kitagawa‡, Kazuko H. Nomura*†§, Katsufumi Dejima*†, Keiko Gengyo-Ando||, Shohei Mitani§||, Kazuyuki Sugahara‡ & Kazuya Nomura*†§

* Department of Biology, Faculty of Sciences 33, Kyushu University, Fukuoka 812-8581, Japan

† SORST and § PRESTO (Japan Science and Technology Corporation) Kawaguchi Center Bldg, 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012, Japan

‡ Department of Biochemistry, Kobe Pharmaceutical University, Higashinada-ku, Kobe 658-8558, Japan

|| Department of Physiology, Tokyo Women's Medical University School of Medicine, Tokyo 162-8666, Japan

Glycosaminoglycans such as heparan sulphate and chondroitin sulphate are extracellular sugar chains involved in intercellular signalling. Disruptions of genes encoding enzymes that mediate glycosaminoglycan biosynthesis have severe consequences in *Drosophila* and mice^{1–5}. Mutations in the *Drosophila* gene *sugarless*, which encodes a UDP-glucose dehydrogenase, impair developmental signalling through the Wnt family member Wingless, and signalling by the fibroblast growth factor and Hedgehog pathways. Heparan sulphate is involved in these pathways^{6–8}, but little is known about the involvement of chondroitin. Under-sulphated and oversulphated chondroitin sulphate chains have been implicated in other biological processes, however, including adhesion of erythrocytes infected with malaria parasite to human placenta and regulation of neural development^{9,10}. To investigate chondroitin functions, we cloned a chondroitin synthase hom-

ologue of *Caenorhabditis elegans* and depleted expression of its product by RNA-mediated interference and deletion mutagenesis. Here we report that blocking chondroitin synthesis results in cytokinesis defects in early embryogenesis. Reversion of cytokinesis is often observed in chondroitin-depleted embryos, and cell division eventually stops, resulting in early embryonic death. Our findings show that chondroitin is required for embryonic cytokinesis and cell division.

Chondroitin glycosaminoglycans (GAGs) are sulphated to varying degrees and, although they are widely distributed throughout the animal kingdom, little is known about the functions of the non-sulphated forms. We previously cloned the genes encoding human chondroitin synthase (ChSy) and chondroitin *N*-acetylgalactosaminyltransferase (ChGn)^{11,12} and identified homologues in *Drosophila melanogaster* and *C. elegans*¹³. Cells from *C. elegans* produce a large amount of non-sulphated chondroitin but not chondroitin sulphate (CS)^{14,15}. To investigate the roles of chondroitin in the development and morphogenesis of this model organism, we depleted the expression of *ChSy* through RNA-mediated interference (RNAi) and monitored embryogenesis of the resulting embryos using four-dimensional (4D) microscopy (multifocal time-lapse image recording).

In humans and *Drosophila*, chondroitin chains are synthesized by ChSy and ChGn, which are required for chain elongation and initiation, respectively. Chain elongation requires GalNAc transferase II (GalNAcT-II) and D-glucuronic acid (GlcA) transferase II (GlcAT-II) activities, whereas chain initiation requires GalNAc transferase I (GalNAcT-I) activity (ref. 5, Fig. 1a and Supplementary Fig. 1a). We thought that *C. elegans* would also possess homologues of the *ChSy* and *ChGn* genes.

A genome database search showed that *C. elegans* has only a *ChSy* orthologue (*T24D1.1*), whereas *D. melanogaster* possesses orthologues of both *ChSy* and *ChGn*. Thus, the *C. elegans* *ChSy* product might function not only in chain elongation, but also in chain initiation in chondroitin synthesis. Two complementary DNA clones corresponding to the predicted *C. elegans* *ChSy* gene *T24D1.1* were isolated by polymerase chain reaction (PCR), with putative amino acid sequences of 715 and 726 residues. The

previously predicted intron 2 region can be translated by alternative splicing, resulting in a larger protein that is 11 amino acids longer than the predicted sequence (Fig. 1b, c). *ChSy* encodes a type II transmembrane protein with two Asp-X-Asp (DXD) motifs (Fig. 1d), which are characteristic of glycosyltransferases found in the Golgi apparatus.

We found that the protein shares 35% identity and 54% similarity with human ChSy and seems to belong to the Fringe-like glycosyltransferase family (Fig. 1d), as it contains the motif Pfam 0243 (ref. 16). *In vitro* expression of the protein showed that it had GalNAcT-I and GalNAcT-II activities. But as we could not detect GlcAT-II activity (Supplementary Fig. 1b, c), another enzyme might carry out this function. Together, these findings indicate that this protein is a chondroitin synthase with GalNAcT-I and GalNAcT-II activities that is involved in chain initiation and elongation of chondroitin in the nematode.

We carried out an immunohistochemical assessment using antibodies against chondroitin. Whole-mount immunohistological staining of adult hermaphrodite worms showed large amounts of chondroitin in the gonads and uterus (Fig. 2a), as well as in oocytes, spermatheca and fertilized eggs. Only weak immunostaining was detected in gonadal arms, which comprise the mitotic zone, the transition zone—where nuclei undergo the transition from the

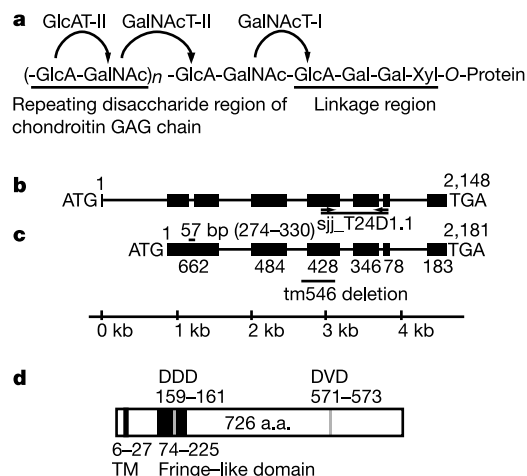


Figure 1 Structure of chondroitin proteoglycan and the nematode *ChSy* gene. **a**, Multiple chondroitin side chains composed of repeating disaccharide units (left horizontal line) are attached to a core protein through a tetrasaccharide sequence (right horizontal line). See text for abbreviations. **b**, The short *ChSy* transcript. **c**, The long *ChSy* transcript, which has an extra 57 base pairs and was used in the RNAi experiments. sjj_T24D1.1 indicates the primer set used for RNAi, which has been shown to have no marked effects²⁸. The position of the deletion in the mutant strain (tm546) is also indicated (horizontal line). **d**, Position of the Asp-Asp-Asp (DDD) and Asp-Val-Asp (DVD) nucleotide-sugar-binding domains in the respective predicted regions of GlcAT and GalNAcT. Filled box, Fringe-like domain; TM, transmembrane domain.

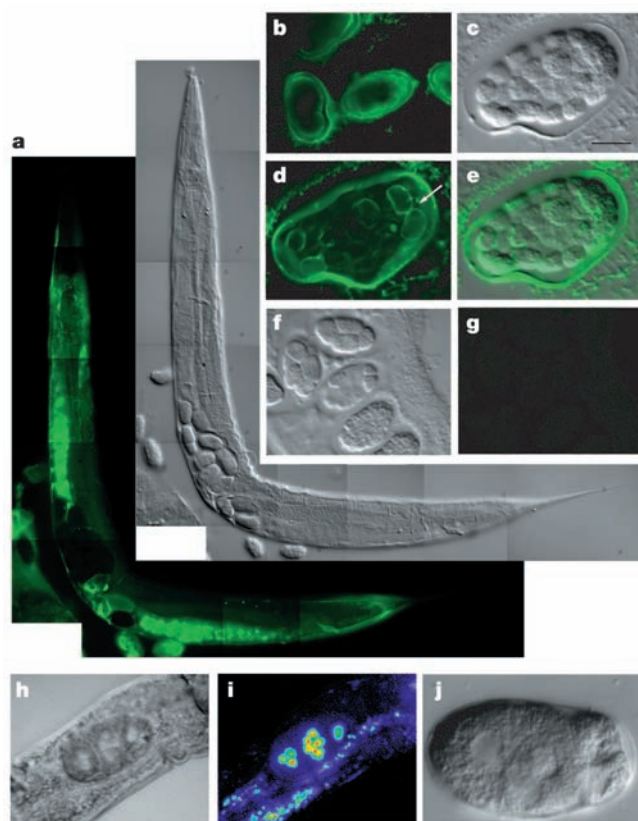


Figure 2 Chondroitin in *C. elegans* and lethal phenotypes in *ChSy*-disrupted worms. **a-g**, Whole-mount immunohistological staining of chondroitin and accompanying DIC microscopic images of worms. Chondroitin was detected with anti-chondroitin antibodies. **a, b**, Freeze-crack method, **c-e**, Finney-Ruvkun method. Large amounts of chondroitin were observed in gonads of adult worms (**a**) and in dividing early embryonic cells and egg shells (**b-e**). The arrow in **d** indicates embryonic cell surface. No staining was observed in control embryos without chondroitinase ABC treatment, indicating the high specificity of the antibody (**f, g**). **h, i**, An RNAi-treated P0 worm. The DIC image (**h**) shows the presence of three cell compartments and the fluorescence image (**i**) shows the presence of multiple nuclei, as indicated by GFP fluorescence. There was no difference in *ChSy*-RNAi phenotypes between N2 and AZ212 strains. **j**, A non-viable egg produced in a *ChSy*^{-/-} deletion mutant worm (tm546). Scale bar, 10 μm.

mitotic cell cycle through early stages of meiotic prophase—and the pachytene zones. Faint staining was detected in the cuticular region. There was abundant chondroitin in fertilized egg shells and at the cell surfaces of cleavage stage embryos (Fig. 2b–g). These results indicate that chondroitin accumulates during oogenesis and may have a role in oogenesis or in early embryogenesis.

Because *C. elegans* contains only one orthologous gene of human *ChSy*, the function of its product can be easily disrupted using RNAi. The full-length *ChSy* cDNA, which corresponds to the larger transcript (Fig. 1c), was used to synthesize double-stranded RNA (dsRNA). The dsRNA was introduced into worms using either the soaking¹⁷ or feeding^{18,19} method to create *ChSy*-RNAi worms. Both methods produced similar, if not identical, phenotypes. Phenotypic abnormalities were first detected in P0 worms 12 h after introducing the dsRNA. Oocytes and fertilized eggs produced in adult P0 hermaphrodites began to die *in utero*. There were no marked effects of RNAi on already hatched P0 worms, including larvae (L1–L3 stage). Soaking L4-stage larvae in the dsRNA solution or feeding

them the bacteria that produce the dsRNA generated morphologically abnormal oocytes and eggs *in utero*, and caused abnormal positioning of eggs in adult hermaphrodites (Fig. 2h, i, and Supplementary Fig. 2a–c; see also Supplementary Movie 1a, b). More than 90% of the F₁ embryos born from the RNAi-treated P0 worms were not viable, and more than 60% of the F₁ adult escapees showed poor gonad formation and laid few eggs. Migration of gonads was distorted, as observed in *ced-2*, *ced-5*, *ced-10* and other cell-migration-abnormal (*mig*) mutant strains (ref. 20 and Supplementary Fig. 2d, e). All F₂ eggs died at early embryonic stages.

To confirm that the RNAi phenotypes were caused by *ChSy* depletion, trimethylpsoralen and ultraviolet-treated deletion libraries of the nematode were screened and the deletion mutant strain tm546 was isolated from pools of mutagenized worms²¹. Because the isolated mutant had a lethal phenotype, it was backcrossed and balanced with *unc-55* (*e402*). The deletion mutant (tm546) lacked 560 base pairs (nucleotides 6741/6742 to 7301/7302 in the cosmid T24D1) of the *ChSy* gene (see Fig. 1c). The mutant strain was maintained as a *ChSy*^{+/–} genotype. A quarter of the F₁ hermaphrodites were expected to have a *ChSy*^{–/–} genotype after self-fertilization. About 50% of the *ChSy*^{–/–} hermaphrodites died and the remaining 50% of this genotype were viable, presumably because of partial rescue by maternal products. The *ChSy*^{–/–} viable F₁ hermaphrodites had abnormal vulvae and produced only a few eggs *in utero*. All of the eggs (F₂) produced from the *ChSy*^{–/–} viable F₁ worms died at various stages during embryonic development (Fig. 2j). The mutant phenotypes observed in the tm546 strain were very similar to those of the *ChSy* RNAi-treated N2 (wild-type) worms. These results indicate that the observed RNAi phenotypes are due to the loss of ChSy activity.

These findings indicated that *ChSy* function is essential in oogenesis or in early embryogenesis. Western blot analysis confirmed that the observed phenotypes were caused by the depletion of chondroitin chains (Fig. 3a). As the worms were fed dsRNA-producing *Escherichia coli*, the chondroitin content decreased at a constant rate to barely detectable amounts after 30 h (Fig. 3b, c). Viability of the RNAi-treated F₁ worms was roughly proportional to the decrease in chondroitin content (Fig. 3d).

RNAi of the *rib-1* and *rib-2* genes encoding putative heparan sulphate (HS)-synthesizing enzymes, however, did not result in similar phenotypes (data not shown). *rib-2* encodes an *N*-acetylglucosaminyltransferase involved in the biosynthetic initiation and elongation of HS²², and *rib-1* encodes a putative HS-GlcAT, although no such enzymatic activity has been shown.

The GAG content of *ChSy*-RNAi F₁ worms was analysed by chondroitinase or heparin lyase digestion, followed by high performance liquid chromatography (HPLC). *ChSy*-RNAi F₁ worms showed a 73% decrease in chondroitin and a 112% increase in HS (Table 1). In addition, the disaccharide composition of the HS from the *ChSy*-RNAi F₁ worms was less sulphated than that of the wild-type worms (Supplementary Table 1). Presumably, the decreased chondroitin synthesis resulted in an increase in the intracellular pool size of UDP-GlcA and UDP-GalNAc/UDP-GlcNAc, which possibly changed the amount and composition of HS. Alternatively, or in addition, HS chains might have been added to core proteins of chondroitin proteoglycans, which lacked chondroitin chains on the common tetrasaccharide linker structure GlcAβ1,3Galβ1,3-Galβ1,4Xylβ1 (ref. 23). Nevertheless, these findings indicate that

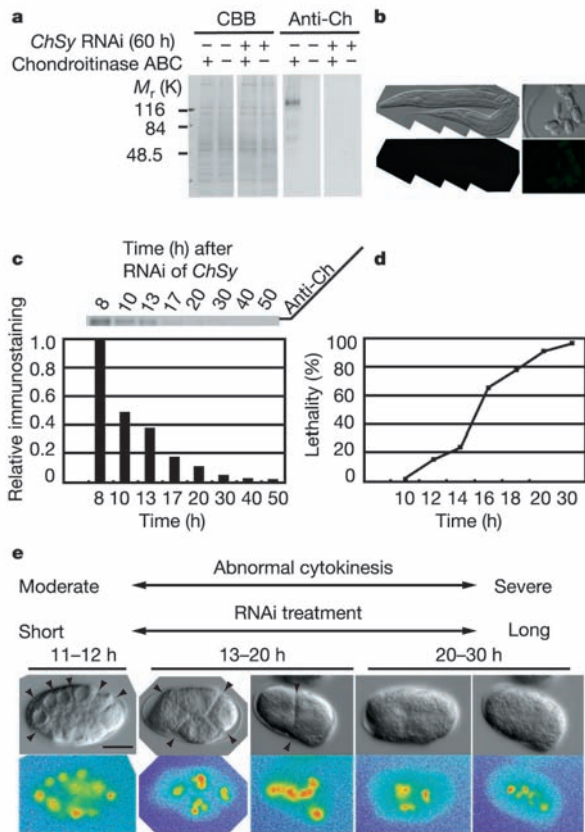


Figure 3 Depletion of chondroitin by RNAi treatment (feeding method) and the effects of chondroitin depletion on embryonic cytokinesis. **a**, Western blot analysis of total protein from *ChSy*-RNAi embryos. The anti-chondroitin antibody recognizes a putative core protein of relative molecular mass 120,000 (*M_r* 120K) that is exposed only after chondroitinase ABC treatment. This protein was detected in control embryos but was absent after 60 h of treatment with *ChSy* RNAi, despite comparable Coomassie brilliant blue (CBB) stained protein patterns. **b**, Absence of chondroitin in the *ChSy* RNAi-treated worms (left) and embryos (right). DIC images (top) and immunofluorescence images with anti-chondroitin antibody (bottom) are shown. **c**, Decrease in chondroitin with increasing RNAi feeding time, as determined by western blot analysis. Results are the mean of three independent experiments. **d**, Lethality of the RNAi treatment. The mean values of five independent experiments are plotted against the dsRNA feeding time. **e**, Short RNAi treatment (11–12 h) induced reversion of cytokinesis, as determined by the increasing and decreasing number of cleavage furrows (arrow heads). Longer RNAi treatment (13–30 h) gave rise to multinucleated embryos without cleavage furrows. Scale bar, 10 μ m.

Table 1 Total amounts of GAG disaccharides		
	N2*	T24D1.1*
Chondroitin	5,874 (100)	1,575 (27)
Heparan sulphate	11.7 (100)	24.8 (212)

*Values are expressed as pmol (%) of disaccharides per mg of dried homogenates of the worms. Numbers in parentheses represent % yields of the chondroitin or heparan disaccharides, taking the respective disaccharides from the N2 worms as 100%. For experimental details, see Methods. T24D1.1, N2 worms treated with T24D1.1 dsRNA.

ChSy is involved in the biosynthesis of chondroitin, but not HS.

We also examined the *ChSy*-RNAi worms by 4D microscopy (Fig. 3e). As dsRNA treatment time increased, chondroitin content decreased (Fig. 3c) and the *ChSy*-RNAi embryos developed severe cell division defects (Fig. 3e). Fertilized eggs laid within 11–12 h of RNAi treatment were not viable and frequently underwent cytokinesis reversal; in addition, embryonic cell numbers oscillated during development, even though nuclear division continued (Fig. 4a, b and Supplementary Movies 2a, b and 3). The reversion of cell division was observed most frequently in embryos that had a moderately decreased content of chondroitin. One such embryo was selectively examined by 4D microscopy immediately after the pseudocleavage stage (Fig. 4a). After the embryo divided into two cells, nuclear division proceeded to form six nuclei. Cytokinesis was not complete, however, leaving only two cell compartments instead of six. The two-cell-stage embryo reverted back to the one-cell-stage embryo and then divided again into two cells, whereas the number of nuclei increased.

Figure 4b shows a typical pattern of reversion of cytokinesis in an RNAi-treated embryo. The two-cell embryo (0 min) divided into a four-cell embryo (4 min), and six nuclei were formed after one nucleic division. The embryo remained in a six-cell stage for 4 min, and then insufficient cytoplasmic division seemed to force the embryonic cell compartments to revert to the four-cell stage (11–19.5 min). Cytoplasmic division continued to form a six-cell-like embryo (27 min). Thereafter, the apparent cell number increased and decreased several times between 31 and 80 min (see Supplementary Movie 3), and eventually the embryo stopped dividing and died. This progression and reversion of embryonic cell division (from two cells, to four cells, to six cells, to four cells, to six cells, and so on) seems to be due to incomplete cytokinesis.

Eggs laid 13–20 h after RNAi treatment produced less chondroitin than those laid 11–12 h after RNAi treatment, and underwent only a few normal cleavages. Consequently, most blastomeres were multinucleated (Fig. 3e). More than 70% of the eggs laid after 20 h or more of continuous RNAi treatment contained very little chondroitin and did not undergo any form of cytokinesis, although nuclear division continued normally. These *ChSy*-RNAi embryos

were therefore single-celled, but contained multiple (10–20) nuclei. The results indicate that chondroitin chains are required for normal cell division and cytokinesis in *C. elegans*.

Glycosaminoglycan chains including HS and CS are synthesized on a unique tetrasaccharide linker structure, GlcA β 1,3Gal β 1,3-Gal β 1,4Xyl β 1 (ref. 23), that is built onto specific serine residues in core proteins of parent proteoglycans (ref. 5 and Fig. 1a). Therefore, inhibition of synthesis of this linker tetrasaccharide should inhibit synthesis of these proteoglycans. RNAi of the *Y110A2AL.14* gene encoding galactosyltransferase II, which is involved in the synthesis of the linker region, produces lethal embryonic phenotypes and several multinucleated embryonic cells¹⁷. Defects in vulval formation and early embryonic death also occur in worms with mutations in the *squashed vulva 8* (*sqv-8*) gene^{24,25}, which encodes a GlcA transferase (GlcAT-I) that is also involved in linker tetrasaccharide synthesis²⁶. The mutant phenotypes observed for the *ChSy*-RNAi worms were similar to those reported for the *sqv-3* and *sqv-8* strains²⁵, in which chondroitin synthesis is affected²⁴. By contrast and as described above, RNAi of the HS-synthesizing enzyme genes *rib-1* and *rib-2* did not produce multinucleated cell death or abnormal cytokinesis (data not shown), suggesting that the abnormal cytokinesis and eventual embryonic cell death observed for these mutants are attributable to the inhibition of the synthesis of chondroitin but not HS.

If chondroitin chains are indispensable for proper embryonic cytokinesis and cell division, digestion of chondroitin chains at the embryonic cell surface might also induce abnormal cell division and reversion of cytokinesis. Treatment of cultured embryonic cells²⁷ with chondroitinase ABC had marked effects on cell division, as observed by 4D microscopy. These cells underwent abnormal cytokinesis, and reversion of cytokinesis was confirmed only in the chondroitin-depleted embryonic cells (Fig. 4c and Supplementary Movie 4).

These results indicate that chondroitin proteoglycans are involved in controlling embryonic cell division and cytokinesis. It is not clear whether they have a structural role, permitting formation of a structural architecture required for cytokinesis, or whether there is an unidentified chondroitin-dependent signalling

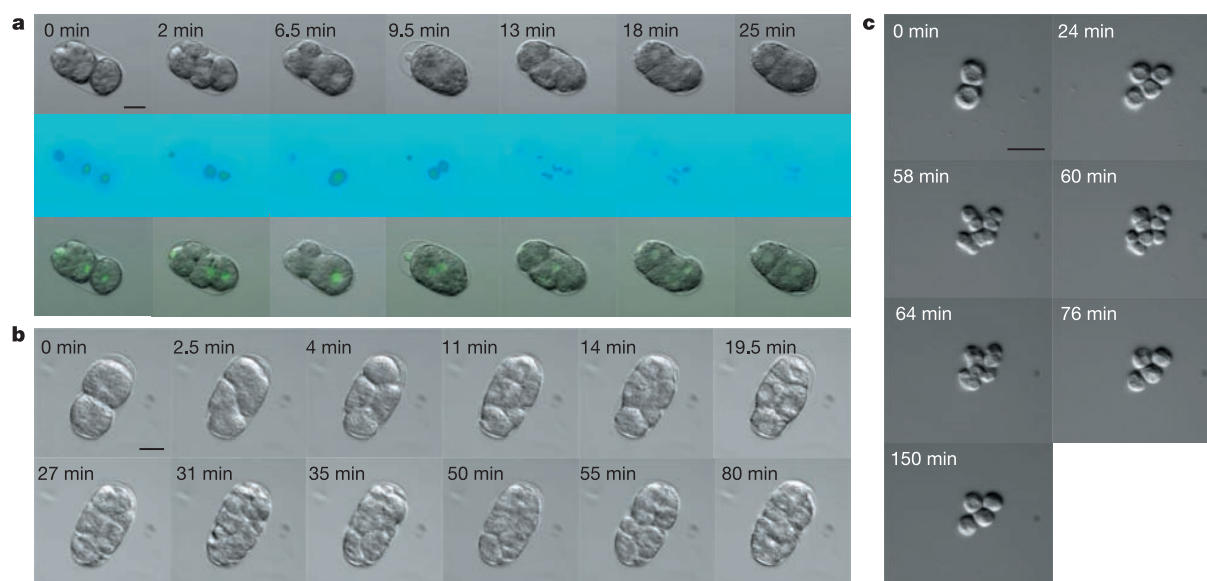


Figure 4 Reversion of cytokinesis in chondroitin-depleted embryos. Fertilized eggs laid 11–12 h after RNAi treatment were analysed by 4D microscopy. Nuclei were visualized by detection of a histone–GFP fusion protein (blue or green). **a**, After the pseudocleavage (not shown), several deep pseudocleavage-like furrows formed in the *ChSy*-RNAi embryo. Over the course of 25 min, cytokinesis failed and the embryo became multinucleated. **b**,

Oscillation of cell division in embryos with reduced chondroitin. **c**, Reversion of cytokinesis in normal embryonic cells treated with chondroitinase ABC. Reversal of cytokinesis was observed in 11 out of 20 dividing embryonic cells treated. No reversal of cytokinesis was observed in 20 non-treated cells (not shown). Scale bar, 10 μ m. See also Supplementary Movies 2–4, which correspond to **a–c**, respectively.

pathway required for completion of cytokinesis. RNAi of the *PAR 2.4* gene, *N*-acetyltransferase gene *T23G11.2*, ubiquitin-related genes *R09B3.4* and *C47B2.4*, and importin gene *F26B1.3* have been also shown to disrupt embryonic cytokinesis (data not shown and ref. 28). Further studies on the relation of these genes to cytokinesis and chondroitin metabolism are needed to clarify the mechanisms of embryonic cell division in *C. elegans*.

The roles of chondroitin proteoglycans in cell division and cytokinesis, as well as associated gene networks, might be most effectively clarified through forward and reverse genetic analysis. Because the *ChSy* deletion mutant has a phenotype that is very similar to the *sqv* mutants, it would be interesting to study the roles of chondroitin proteoglycans and chondroitin-degrading enzymes in vulval formation and somatic cell fusion. So far, there are no reports on the roles of proteoglycans in cytokinesis, although chondroitin or CS proteoglycans might have similar roles in cytokinesis and cell division in higher organisms. In this context, it is noteworthy that *C. elegans* does not seem to synthesize hyaluronan¹⁵. This proteoglycan might have roles in higher organisms similar to those of chondroitin observed in this study. In fact, hyaluronan is involved in cell division of human fibroblasts²⁹. To clarify the roles of proteoglycans in cytokinesis, further studies on the relationship of chondroitin proteoglycans with cytoplasmic actin, tubulin filaments and cell cycle-regulating proteins are needed. □

Methods

Strains

All the nematode and *E. coli* strains used were from the Caenorhabditis Genetics Centre.

Analysis of GAGs

The GAG fraction was analysed as described¹⁵, using the dried homogenates of wild-type or RNAi-treated nematodes (dry weight, 56 or 70 mg, respectively). We released GAG chains from the proteoglycan core proteins by sodium borohydride treatment. It should be noted that the amount of HS in *C. elegans* was so small¹⁵ that 100 µg of shark cartilage chondroitin 6-O-sulphate (Seikagaku), which contains a negligible proportion of non-sulphated disaccharides, was added as a carrier after the borohydride treatment but before the purification steps. The purified GAG fraction was digested with chondroitinase ABC or a mixture of heparin lyases I and II, and then the digests were derivatized with 2-aminobenzamide and analysed by HPLC to determine the disaccharide compositions.

Western blotting

We collected the worms by centrifugation after washing with M9 buffer¹⁷. The *C. elegans* pellet (10 mg) was sonicated with a GE-70 ultrasonic processor (Branson Ultrasonics) in ice-cold 20 mM sodium phosphate buffer (pH 7.2) and immediately heated at 90°C for 20 min to inactivate proteases. After centrifugation at 3,000g for 5 min, the supernatant fluid was used for protein determination and western blot analysis, followed by chondroitinase ABC digestion. We resolved the proteins on 10% SDS polyacrylamide gels, transferred them to nitrocellulose membranes and incubated them with an anti-chondroitin antibody (anti-proteoglycan ΔDi-OS antibody 1-B-5; Seikagaku) for 3 h. This antibody recognizes the unsaturated hexuronic acid (HexA)-GalNAc epitope generated by chondroitinase treatment. The anti-chondroitin antibody was diluted 1:2,000 with 25 mM Tris-buffered saline (TBS) before use. After washing, the bound antibody was detected with a F(ab')₂ against mouse IgG conjugated to alkaline phosphatase (EY Laboratory) diluted 1:5,000 with NBT/BCIP. We stopped the reaction with TBS containing 20 mM EDTA and scanned the stained bands with an optical scanner. The intensity of the staining was quantified using Scion Image Software (Scion).

Immunofluorescence microscopy

The worms were washed once with M9 buffer, fixed and permeabilized by the Finney–Ruvkun protocol or a freeze-crack protocol³⁰. After blocking with TBS containing 0.1% Tween 80, worms were treated with 0.05 U µl⁻¹ chondroitinase ABC for 1 h and incubated with the anti-chondroitin antibody diluted 1:2,000 with TBS at 20°C for 2–4 h. After being washed, the worms were incubated with an antibody against mouse IgG conjugated to fluorescein isothiocyanate for 2 h at 20°C. After being washed in TBS, each specimen was covered with a mounting solution containing 5% α,1,4-diazabicyclo(2,2,2)octane (Dabco) in 95% glycerol and analysed by fluorescence microscopy (DMRXA full automatic microscope; Leica). We recorded the images with a digital CCD (charge-coupled device) camera (Micro MAX-1300; Roper) and processed them using the MetaMorph software (version 4.6; Universal Imaging).

Four-dimensional microscopy

The main hardware for the 4D microscopy system comprised an Optiplex GX1 Computer (Dell), a DMRXA microscope (Leica) with differential interference contrast (DIC) and fluorescence optics, and the CCD camera described above. A shutter controller (VMM-D3; Uniblitz) and a microscope focus controller (Leica) were attached to the

microscope to control focus and illumination automatically. The mother worms treated with dsRNA were dissected in M9 buffer, the eggs or young embryos were collected and put on an eight-well printed microscope slide glass (Matsunami Glass), and the coverslip was sealed with white petroleum jelly. We recorded the developing embryos at 20°C for 8–20 focal planes, at a distance of 1–2 µm between adjacent focal planes and at intervals of 30–120 s for 1–8 h under DIC or fluorescence (green fluorescent protein, GFP) illumination. GFP-positive nuclei were visualized by using GFP-tagged histone-expressing worms (Caenorhabditis Genetics Center strain AZ212) as *ChSy*-RNAi hosts. The data were acquired with the CCD camera and movies were made from the raw data using MetaMorph software (Universal Imaging).

RNAi by soaking or feeding

The soaking protocol was based on described methods¹⁷. The cDNA for *T24D1.1* (*ChSy*) was amplified by PCR of total cDNA from N2 worms, and the other gene fragments were amplified by PCR of cDNA from N2 worms or from the Expressed Sequence Tag project¹⁷. Primers for the *T24D1.1* cDNA were 5'-CGGGATCCAGAAATGCCGTCTCCCGAG-3' and 5'-CGGGATCCGGTGGAAAACCGCAAGG-3'. Each cDNA fragment was cloned into a pBlueScriptII SK phagemid vector or a pCMV-Script vector (Stratagene). Cloned fragments were re-amplified with a T3 (5'-AATTAAACCTCACTAAAGGGAAC-3') and a T7 (5'-GTAATACGACTCACTATAGGGCGA-3') primer. The PCR products (50–150 µl) were purified and used for *in vitro* transcription with T3 and T7 RNA polymerases (Stratagene). L4-stage hermaphrodites were washed with M9 buffer and soaked in 4 µl of the *ChSy* dsRNA solution (0.5–1 µg µl⁻¹) at 20°C for 24 h.

The protocol for RNAi-by-feeding was based on described methods^{18,19}. Each cDNA fragment was cloned into the L4440 (pPD129.36) vector and the cloned plasmids were transformed into *E. coli* HT115 (DE3). A single colony of HT115 (DE3) containing the plasmid was grown in LB culture for 8 h and seeded onto NGM agar plates (80 µl per plate), which were incubated at 37°C overnight. We then added 10 mM isopropyl-β-D-thiogalactoside (IPTG, 80 µl per plate) and cultivated them for 4 h to induce the expression of dsRNA. Mixed-stage hermaphrodites were sown on these plates and dsRNA was introduced into the nematode by feeding. After dsRNA introduction, the worms were washed and collected by centrifugation as described above.

C. elegans embryonic cell cultures

Cells of *C. elegans* embryos were cultivated as described²⁷. In brief, the egg shells of bleach-collected eggs were removed by treatment with 1 U ml⁻¹ chitinase (Sigma) at 20°C for 8 min, the cell suspension was passed through a 23-gauge needle, and the cells were washed with Leibovitz's L-15 cell culture medium (345 mOsm adjusted with sucrose; Gibco-BRL) containing 10% fetal bovine serum (Cell Culture Technologies). The embryonic cells were further filtered through a sterile 5-µm Durapore syringe filter (Millipore) and plated on eight-well printed microscope slides, which had been coated with 1 mg ml⁻¹ peanut lectin agglutinin (Sigma).

For chondroitinase ABC digestion, we cultivated embryonic cells at 20°C in the optimized L-15 medium described above and added chondroitinase ABC to the culture medium at a final concentration of 0.05 U µl⁻¹.

Received 28 January; accepted 21 March 2003; doi:10.1038/nature01635.

1. Forsberg, E. & Kjellén, L. Heparan sulfate: lessons from knockout mice. *J. Clin. Invest.* **108**, 175–180 (2001).
2. Iozzo, R. V. Heparan sulfate proteoglycans: intricate molecules with intriguing functions. *J. Clin. Invest.* **108**, 165–167 (2001).
3. Lander, A. D. & Selleck, S. B. The elusive functions of proteoglycans: *in vivo* veritas. *J. Cell Biol.* **148**, 227–232 (2000).
4. Perrimon, N. & Bernfield, M. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* **404**, 725–728 (2000).
5. Sugahara, K. & Kitagawa, H. Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr. Opin. Struct. Biol.* **10**, 518–527 (2000).
6. Hacker, U., Lin, X. & Perrimon, N. The *Drosophila* sugarless gene modulates Wingless signaling and encodes an enzyme involved in polysaccharide biosynthesis. *Development* **124**, 3565–3573 (1997).
7. Lin, X., Buff, E. M. & Perrimon, N. Heparan sulfate proteoglycans are essential for FGF receptor signaling during *Drosophila* embryonic development. *Development* **126**, 3715–3723 (1999).
8. Tsuda, M. *et al.* The cell-surface proteoglycan Dally regulates Wingless signalling in *Drosophila*. *Nature* **400**, 276–280 (1999).
9. Bradbury, E. J. *et al.* Chondroitinase ABC promotes functional recovery after spinal cord injury. *Nature* **416**, 636–640 (2002).
10. Sugahara, K. & Yamada, S. Structure and function of oversulfated chondroitin sulfate variants: unique sulfation patterns and neuroregulatory activities. *Trends Glycosci. Glycotechnol.* **12**, 321–349 (2000).
11. Uyama, T., Kitagawa, H., Tamura, J. & Sugahara, K. Molecular cloning and expression of human chondroitin *N*-acetylgalactosaminyltransferase: the key enzyme for chain initiation and elongation of chondroitin/dermatan sulfate on the protein linkage region tetrasaccharide shared by heparin/heparan sulfate. *J. Biol. Chem.* **277**, 8841–8846 (2002).
12. Uyama, T. *et al.* Molecular cloning and expression of a second chondroitin *N*-acetylgalactosaminyltransferase involved in the biosynthetic initiation and elongation of chondroitin/dermatan sulfate. *J. Biol. Chem.* **278**, 3072–3078 (2003).
13. Kitagawa, H., Uyama, T. & Sugahara, K. Molecular cloning and expression of a human chondroitin synthase. *J. Biol. Chem.* **276**, 38721–38726 (2001).
14. Toyoda, H., Kinoshita-Toyoda, A. & Selleck, S. B. Structural analysis of glycosaminoglycans in *Drosophila* and *Caenorhabditis elegans* and demonstration that *tout-velu*, a *Drosophila* gene related to EXT tumour suppressors, affects heparan sulfate *in vivo*. *J. Biol. Chem.* **275**, 2269–2275 (2000).
15. Yamada, S. *et al.* Demonstration of glycosaminoglycans in *Caenorhabditis elegans*. *FEBS Lett.* **459**, 327–331 (1999).
16. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **30**, 276–280 (2002).

17. Maeda, I., Kohara, Y., Yamamoto, M. & Sugimoto, A. Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Curr. Biol.* **11**, 171–176 (2001).
18. Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* **2**, research0002.1–0002.10 [online] (<http://genomebiology.com/2000/2/11/research/0002>) (2001).
19. Timmons, L., Court, D. L. & Fire, A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263**, 103–112 (2001).
20. Reddien, P. W. & Horvitz, H. R. CED-2/CrkII and CED-10/Rac control phagocytosis and cell migration in *Caenorhabditis elegans*. *Nature Cell Biol.* **2**, 131–136 (2000).
21. Gengyo-Ando, K. & Mitani, S. Characterization of mutations induced by ethyl methanesulfonate, UV, and trimethylpsoralen in the nematode *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **269**, 64–69 (2000).
22. Kitagawa, H. *et al.* *rib-2*, a *Caenorhabditis elegans* homolog of the human tumour suppressor *EXT* genes encodes a novel α 1,4-*N*-acetylglucosaminyltransferase involved in the biosynthetic initiation and elongation of heparan sulfate. *J. Biol. Chem.* **276**, 4834–4838 (2001).
23. Lindahl, U. & Rodén, L. in *Glycoproteins* (ed. Gottschalk, A.) 491–517 (Elsevier, New York, 1972).
24. Bulik, D. A. *et al.* *sqv-3*, *-7*, and *-8*, a set of genes affecting morphogenesis in *Caenorhabditis elegans*, encode enzymes required for glycosaminoglycan biosynthesis. *Proc. Natl Acad. Sci. USA* **97**, 10838–10843 (2000).
25. Herman, T., Hartwig, E. & Horvitz, H. R. *sqv* mutants of *Caenorhabditis elegans* are defective in vulval epithelial invagination. *Proc. Natl Acad. Sci. USA* **96**, 968–973 (1999).
26. Kitagawa, H. *et al.* Molecular cloning and expression of glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **273**, 6615–6618 (1998).
27. Christensen, M. & Strange, K. Developmental regulation of a novel outwardly rectifying mechanosensitive anion channel in *Caenorhabditis elegans*. *J. Biol. Chem.* **276**, 45024–45030 (2001).
28. Kamath, R. S. *et al.* Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237 (2003).
29. Brecht, M., Mayer, U., Schlosser, E. & Prehm, P. Increased hyaluronate synthesis is required for fibroblast detachment and mitosis. *Biochem. J.* **239**, 445–450 (1986).
30. Miller, D. M. & Shakes, D. C. *Caenorhabditis elegans*. in *Modern Biological Analysis of an Organism* (eds Epstein, H. F. & Shakes, D. C.) 365–394 (Academic, London, 1995).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank T. Stiernagle and the *Caenorhabditis* Genetics Center for all worms and *E. coli* strains, and Y. Kohara for the *yk* clones. K.N. was supported by PRESTO and SORST of the Japan Science and Technology Corporation (JST). S.M. and K.H.N. were supported partly by Sasakawa Scientific Research Grant (JSS). The work at Kobe Pharmaceutical University was supported in part by a Science Research Promotion Fund from the Japan Private School Promotion Foundation, and by Grants-in-Aid for Scientific Research C (to H.K.) and Scientific Research B (to K.S.) from the Ministry of Education, Science, Sports and Culture of Japan.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to K.N. (knomusc@mbbox.nc.kyushu-u.ac.jp) or K.S. (k-sugar@kobepharm-u.ac.jp). Sequences of the longer ChSy and the shorter ChSy have been deposited in the DNA Data Bank of Japan under accession numbers AB088397 and AB088398, respectively.

Wnt proteins are lipid-modified and can act as stem cell growth factors

Karl Willert*, Jeffrey D. Brown†, Esther Danenberg*, Andrew W. Duncan†, Irving L. Weissman‡, Tannishtha Reya†, John R. Yates III§ & Roel Nusse*

* Howard Hughes Medical Institute and Department of Developmental Biology, Stanford University School of Medicine, Stanford, California 94305, USA

† Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina 27710, USA

‡ Departments of Pathology and Developmental Biology, Stanford University School of Medicine, Stanford, California 94305, USA

§ Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037, USA

Wnt signalling is involved in numerous events in animal development¹, including the proliferation of stem cells² and the specification of the neural crest³. Wnt proteins are potentially important reagents in expanding specific cell types, but in contrast to other developmental signalling molecules such as

hedgehog proteins and the bone morphogenetic proteins, Wnt proteins have never been isolated in an active form. Although Wnt proteins are secreted from cells^{4–7}, secretion is usually inefficient⁸ and previous attempts to characterize Wnt proteins have been hampered by their high degree of insolubility. Here we have isolated active Wnt molecules, including the product of the mouse *Wnt3a* gene. By mass spectrometry, we found the proteins to be palmitoylated on a conserved cysteine. Enzymatic removal of the palmitate or site-directed and natural mutations of the modified cysteine result in loss of activity, and indicate that the lipid is important for signalling. The purified Wnt3a protein induces self-renewal of haematopoietic stem cells, signifying its potential use in tissue engineering.

We expressed several Wnt genes, including *Wnt3a* (ref. 9), in a variety of cell lines and generated antibodies to monitor Wnt protein secretion into the medium. For purification purposes, we selected clones of cells secreting the highest amounts of protein (200 ng ml^{−1} for Wnt3a from mouse L (L-M[TK[−]]) cells). We tested the activity of Wnt3a by assaying its ability to stabilize cytosolic β -catenin, a known target and signal transduction component of Wnt signalling¹⁰. Mouse L cells accumulate high levels of β -catenin protein after a 2-h incubation with Wnt3a-conditioned medium (Fig. 1b, top panel; see also ref. 11).

Initial characterization of secreted Wnt3a indicated that it is hydrophobic (see below), therefore we designed a purification protocol that starts with chromatography on blue (Cibacron blue 3GA) Sepharose in the presence of the detergent CHAPS. Under these conditions, Wnt3a binds with high selectivity to the resin and can be eluted in a relatively pure form by increasing ionic strength (Fig. 1a and Table 1). Approximately 60% of added Wnt3a is recovered in this step with nearly 2,500-fold enrichment. We then separated Wnt-containing fractions by size exclusion chromatography on a Superdex 200 column, and finally by cation exchange on heparin (Table 1). These steps yielded fractions of Wnt3a that were greater than 95% pure as assessed by Coomassie staining (Fig. 1a). Through size exclusion chromatography, we determined that active Wnt3a is monomeric (not shown).

We have applied successfully similar purification methods to a variety of other Wnt proteins, including *Drosophila* Wnt8 (Fig. 1a), mouse Wnt5a and *Drosophila* Wingless (not shown).

Throughout the purification, we measured the ability of Wnt3a to stabilize β -catenin in L cells. The final purified product exhibited no loss in activity compared to the original starting material (Fig. 1b). The purified Wnt3a protein retains the range of activities expected for a Wnt protein. For example, we tested the effect of Wnt3a protein on *Xenopus* animal cap explants and found that two known target genes, *siamois* and *Xnr3* (refs 12, 13), are induced by Wnt3a (Fig. 1c). As a further assay for Wnt activity, we used C57MG cells, a line derived from the mouse mammary gland that can be morphologically transformed by *Wnt* gene expression⁸. Purified Wnt3a promotes the morphological transformation of these cells (Fig. 1d) similar to that of *Wnt* gene transfection. Furthermore, the protein can induce expression of known transcriptional Wnt targets including *MSX1*, cyclin D1 and *MYC* in human teratocarcinoma cells (data not shown).

All purification steps required the presence of detergent to maintain solubility and activity, suggesting that Wnt proteins are hydrophobic. We used the two-phase separation property of the detergent Triton X-114 (ref. 14) to test this. Most Wnt3a partitioned to the detergent phase (Fig. 2a), a behaviour characteristic of highly hydrophobic proteins such as integral membrane proteins. As the primary amino acid sequence of secreted Wnt does not contain long stretches of hydrophobic residues, we used metabolic labelling to test whether Wnt is post-translationally modified by lipid attachment. We found that the protein is labelled with tritiated palmitate (Fig. 2b).

Evidence for the functional importance of the lipid modification

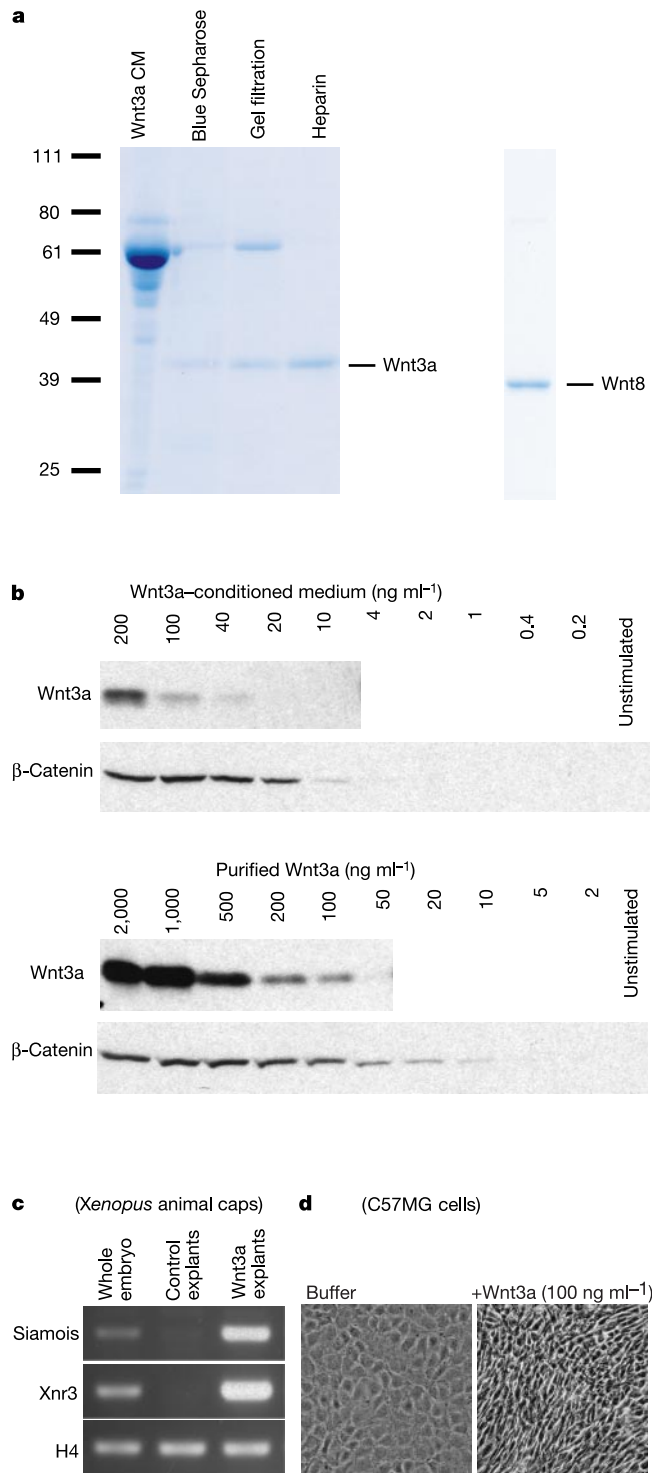


Figure 1 Wnt3a and *Drosophila* Wnt8 purification. **a**, Coomassie staining of an SDS polyacrylamide gel containing fractions from all steps of the purification reveals the enrichment of the Wnt3a protein. Also shown is the final *Drosophila* Wnt8 fraction, purified using the same protocol. Size markers are in kilodaltons. **b**, Wnt3a stabilizes the β -catenin protein. Wnt3a-conditioned medium (200 ng ml⁻¹) and purified Wnt3a (100 μ g ml⁻¹) was diluted as indicated in medium containing 10% FBS and detected by western blot. L cells were stimulated for 2 h. **c**, Wnt3a induces expression of *siamois* and *Xnr3* in animal cap explants of *Xenopus* embryos. Animal cap explants were incubated with 100 ng ml⁻¹ Wnt3a and analysed by polymerase chain reaction with reverse transcription for expression of the direct targets *Xnr3* and *siamois*. **d**, Wnt3a induces the morphological transformation of C57MG cells. C57MG cells were treated with or without 100 ng ml⁻¹ Wnt3a for 2 days in serum-containing medium and then an additional 2 days in serum-free medium.

came from treatment of Wnt3a with acyl-protein thioesterase-1 (APT-1), an enzyme that removes palmitate from G proteins and other thioacyl protein substrates¹⁵. This treatment shifts Wnt3a to the water phase in the Triton X-114 phase separation experiment (Fig. 2c), suggesting that APT-1 removes a thioester-linked lipid, such as palmitate. APT-1 also blocks the ability of Wnt3a to stabilize β -catenin (Fig. 2c).

To map the lipid attachment site on the Wnt polypeptide we subjected proteolytic peptide fragments of both Wnt3a and *Drosophila* Wnt8 to liquid chromatography tandem mass spectrometry, which identifies molecular masses of the ionized peptides and obtains primary amino acid sequence information through collision-induced fragmentation. In both proteins we identified ions whose masses were consistent with the addition of 238 daltons (the mass of palmitate is 256 daltons accounting for the loss of water in the formation of a thioester linkage) and which produced fragmentation data consistent with a peptide containing a conserved cysteine modified by palmitate (C77 in Wnt3a and C51 in *Drosophila* Wnt8; underlined in Fig. 2d). This cysteine is absolutely conserved among all Wnt family members (bold in Fig. 2d); it is the most amino-terminally conserved cysteine of the Wnt family (<http://www.stanford.edu/~rnusse/genealigns/manywnts.html>).

To test for the requirement of C77 in cell culture, we mutated it to alanine in Wnt3a and expressed the mutant protein (Wnt3a(C77A), Fig. 2e) in 293 and in L cells. The mutant Wnt3a protein was secreted at levels similar to that of the wild-type protein. This indicated that the mutation, unlike many other cysteine mutations in Wnt proteins¹⁶, does not interfere with the folding of the protein. However, when the Wnt3a(C77A) protein was subjected to the Triton X-114 phase separation test, it partitioned in the water phase, indicating that it had lost its hydrophobic character (Fig. 2a). In a β -catenin assay on L cells, Wnt3a(C77A) was not active over a range of concentrations tested (Fig. 2e, left). In a transfection assay on 293 cells however, there was a noticeable increase in the intracellular levels of β -catenin, demonstrating that the Wnt3a(C77A) mutant retains some activity when expressed at high levels in an autocrine manner (Fig. 2e, right).

Notably, a natural loss-of-function allele of the *Caenorhabditis elegans* *egl-20* gene (*egl-20(N585)*; ref. 17) contains a serine replacing the cysteine corresponding to C77 (Fig. 2d). Moreover, in a survey of *wingless* (*wg*) alleles in *Drosophila*, we found that the *wg*^{S21} allele¹⁸ contains a tyrosine instead of that same cysteine (Fig. 2d). Thus, our data are consistent with the lipid modification being important for Wnt signalling activity. At the moment, we cannot exclude the possibility that Wnt proteins carry other modifications beyond palmitoylation and N-linked glycosylation; nor can we rule out that different forms of Wnt proteins (that is, cell-bound) are palmitoylated at other sites.

Next we investigated whether Wnt3a can be used as a reagent to control cell fate in a well-characterized stem cell system, through application of the isolated protein to purified haematopoietic stem cells (HSCs)¹⁹. Single HSCs responded well to the Wnt3a protein in the presence of limiting doses of steel factor (SLF). Over a period of 7 days, the frequency of cells proliferating was 5.8-fold greater compared with control conditions (Fig. 3a, b). Most of the cells (82%) were undifferentiated, as they did not express markers for differentiated lineages. Thirty per cent of the lineage-negative cells expressed c-Kit and Sca-1, consistent with an HSC phenotype, whereas 64% were at the stage of myeloid progenitors (c-Kit⁺, Sca-1⁺; Fig. 3c, d). In contrast, incubation of HSCs with unfractionated Wnt3a-conditioned medium, in which Wnt3a itself is present at a similar concentration, resulted in a significant fraction (86%) of the cells expressing markers specific for differentiated lineages (Fig. 3c). This suggests that conditioned medium contains factors not present in purified Wnt3a that promote differentiation, underscoring the importance of having purified Wnt proteins available for the purpose of maintaining the self-renewing fate of HSCs.

Table 1 Purification table

	Volume (ml)	Protein concentration	Total protein (mg)	Wnt3a concentration	Wnt3a (μ g)
Wnt3a CM	2,000	4.46*	8,920	200‡	400
Blue Sepharose	60	36.0†	2.16	4†	240
Gel filtration	36	17.1†	0.615	5†	180
Heparin cation exchange	1.15	104†	0.120	100†	115

The concentration of Wnt3a protein in the conditioned medium (CM) was determined by comparing its signal intensity on a Wnt3a immunoblot to that of a serial dilution of a known amount of purified Wnt3a protein.
*Concentration in mg ml^{-1} .
†Concentration in $\mu\text{g ml}^{-1}$.
‡Concentration in ng ml^{-1} .

To determine whether the cells that proliferated in response to Wnt3a truly maintained HSC activity, we carried out a transplantation analysis. Single HSCs were plated in Terasaki plates and treated with Wnt3a or control medium for a period of 6 days. In previous experiments we showed that culturing cells with SLF alone (our control conditions) while inducing proliferation does not induce self-renewal *in vitro*¹⁹. Each well containing cells that responded to Wnt3a from a single cell was separately injected into lethally irradiated mice, and analysed after 6 weeks of reconstitution (Fig. 3e). If no self-renewal had occurred, only 10% of the mice would be expected to be reconstituted successfully (see Fig. 3 legend). In contrast 100% of the transplanted mice contained donor-derived cells (Fig. 3f), suggesting that HSCs had undergone self-renewal in response to purified Wnt3a.

We have established methods to purify significant quantities of pure and active Wnt proteins, which can be used for self-renewal of

HSCs and potentially other stem cells. We found that Wnt proteins are unexpectedly hydrophobic and are post-translationally modified by palmitoylation, a property that explains the poor solubility of the proteins. It is interesting to note that the protein products of the *Drosophila porcupine* and *C. elegans mom-1* genes^{20,21} have homology with acyl transferases and may catalyse Wnt acylation²². Moreover, the Porcupine protein can bind to a domain in Wingless encompassing the acylation site²³. *porcupine* and *mom-1* have phenotypes similar to *Wnt* alleles and are required in Wnt-producing cells, indicating that the lipid is an integral part of signalling activity. However, overexpression of Wingless in the *Drosophila* embryo can overcome the absence of *porcupine*²⁴, just as high expression of Wnt3a(C77A) can lead to a modest increase in β -catenin (Fig. 2d). This suggests that the lipid functions to increase the local concentration of Wnt on membranes, and that its absence can be overcome by high levels of expression. Although palmitoylation of secreted proteins seems unusual, there is an intriguing

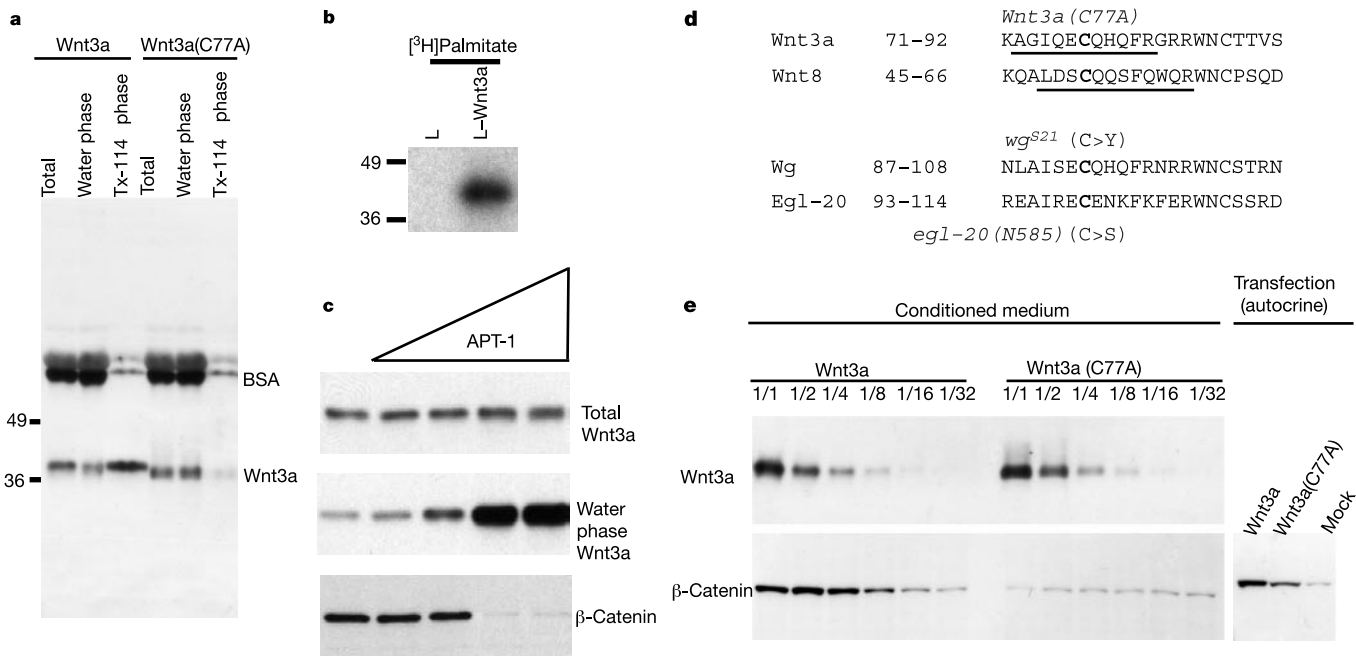


Figure 2 Wnt proteins are palmitoylated on an essential cysteine. **a**, Triton X-114 phase separation (western blot). Most wild-type Wnt3a separates to the Triton X-114 phase, indicating that it is hydrophobic, but the Wnt3a(C77A) mutant (see **d**) partitions mostly to the water phase. BSA from serum partitions to the water phase and serves as an internal control. **b**, *In vivo* labelling of Wnt3a protein with tritiated palmitate. Wnt3a was partially purified from conditioned medium of cells labelled with tritiated palmitate for 5 h. **c**, APT-1 treatment of Wnt3a (western blot). Treatment of Wnt3a with increasing amounts of APT-1 shifts the Wnt3a protein from the Triton X-114 phase (data not shown) to the water phase (middle panel) and abolishes its activity in the β -catenin stabilization assay. **d**, Mass spectrometry maps the palmitate modification to a cysteine (bold) in Wnt3a (C77) and in *Drosophila* Wnt8 (C51). Underlined sequence corresponds to the peptide identified in the

spectra as being modified. The cysteine is conserved in all known Wnt proteins. A site-directed mutant (Wnt3a(C77A)) was made and used in **a** and **e**. The *Drosophila* *wg*^{S21} (ref. 18) allele has a mutation converting the cysteine into a tyrosine and the *egl-20(N585)* allele in *C. elegans* has a serine instead of the cysteine¹⁷. These are both loss-of-function alleles. **e**, The Wnt3a(C77A) mutant protein is secreted from 293 cells at levels similar to wild type, but is not active in increasing β -catenin in target L cells over a range of concentrations tested (western blot). However, the 293 cells transfected with the Wnt3a(C77A) expression construct show a modest increase in β -catenin, indicating that high levels of the mutant can activate Wnt signalling. The Wnt3a(C77A) and wild-type transfected cells express equal levels of Wnt protein.

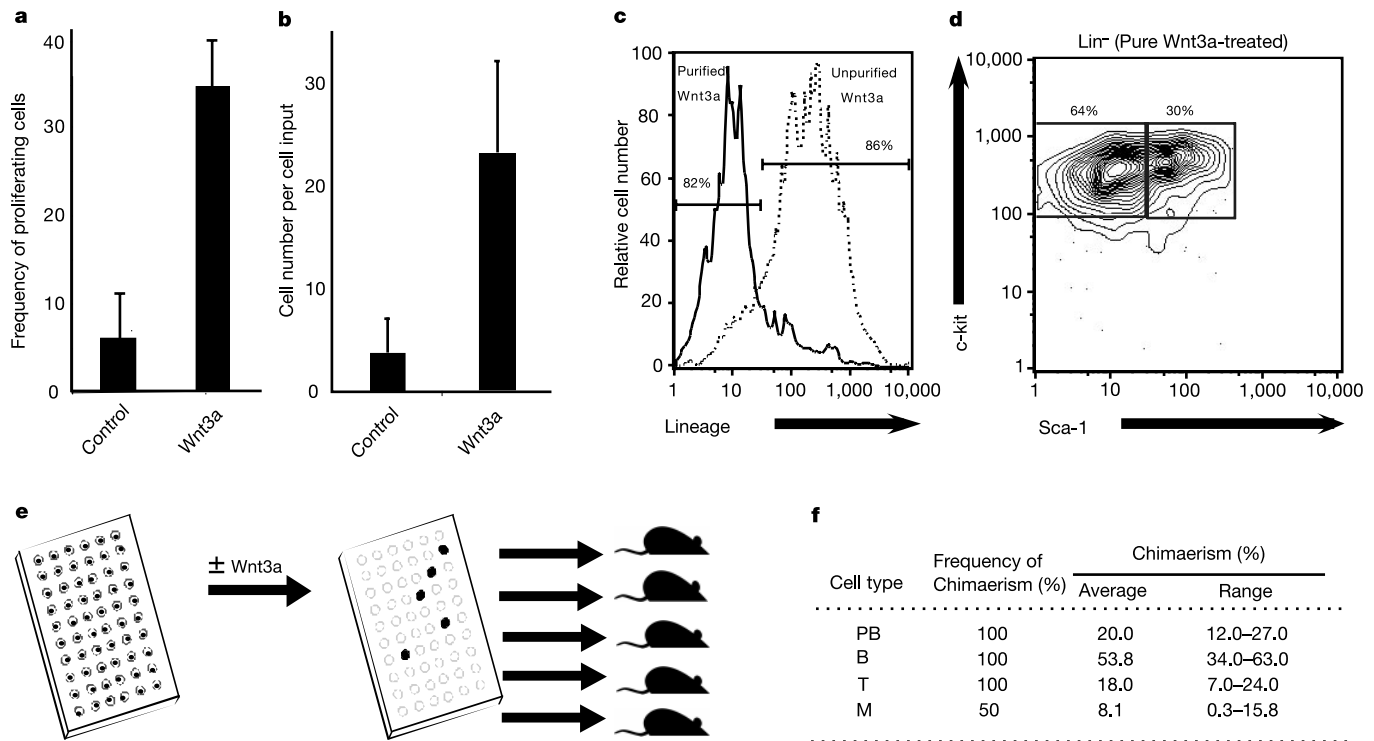


Figure 3 HSCs maintain self-renewing fate with reduced differentiation in response to purified Wnt3a. Purified mouse bone marrow (BM) HSCs (c-Kit⁺, Sca-1⁺, Thy-1.1^{lo}, Lin⁻) from Bcl2 transgenic mice¹⁹ were sorted by FACS and plated as single cells into 60-well Terasaki plates. Cells were incubated in X-vivo15 (Bio Whittaker) containing either purified Wnt3a (about 100 ng ml⁻¹) plus limiting amounts of SLF (7.5 ng ml⁻¹) or SLF (7.5 ng ml⁻¹) alone, as a control. **a**, Cell growth was monitored over 7 days in culture, and shown as the frequency of responding cells. **b**, Total cell growth. Cells responded to Wnt3a by proliferating >100-fold (from 1 cell to at least 100 cells) and the total number of cells generated was sixfold greater in the presence of Wnt3a than control conditions. Results are representative of four independent experiments. **c**, To determine phenotypic characteristics, cells were plated in 96-well plates and incubated in the presence of purified or unpurified Wnt3a. After 7 days in culture, most cells treated with purified Wnt3a (at 100 ng ml⁻¹) were negative for lineage markers (solid line) whereas most treated with unpurified Wnt3a (200 ng ml⁻¹ in the medium; Table 1) strongly upregulated lineage markers (dotted line). **d**, FACS analysis of purified Wnt3a-treated cells. The

lineage-negative population is distributed into c-Kit⁺ and Sca-1⁺ HSCs and c-Kit⁺ and Sca-1⁻ myeloid progenitors. **e**, Purified mouse BM HSCs were plated singly into 60-well Terasaki plates and treated with Wnt3a for 6 days, then all cells generated from the single cell were transplanted individually into lethally irradiated mice along with 300,000 rescuing BM cells. **f**, Peripheral blood (PB) from each transplanted mouse was analysed after 6 weeks for reconstitution along both lymphoid (B and T) and myeloid (M) lineages. On the basis of the reconstitution efficiency of single transplanted HSCs, 1 of 10 (10%) resting HSCs and probably 1 of 50 (2%) cycling HSCs reconstitute. A 50% reconstitution rate suggests at least a fivefold and probably a 15–25-fold expansion in HSCs per transplant. Fivefold expansion is probably an underestimate as HSCs transplanted in low numbers lead to low and variable reconstitution. But our finding that Wnt3a-treated HSCs on transplantation lead to an average chimaerism of 20% (range 12–27%) in the context of a competitive reconstitution suggests a greater than fivefold expansion of functional HSCs.

parallel between Wnt and hedgehog signalling, as the hedgehog protein is also palmitoylated²⁵. □

Methods

Purification of Wnt3a

Mouse L cells (American Type Culture Collection (ATCC) CRL-2648) were cultured in DMEM medium, 10% fetal bovine serum (FBS) and antibiotics. These cells were stably transfected with a vector containing the Wnt3a complementary DNA under the control of the PGK promoter, and G418-resistant clones were selected and screened for production of Wnt3a protein (ATCC CRL-2647). *Drosophila* S2 cells were used to produce the *Drosophila* Wnt8 protein, which was expressed from a heat-shock promoter²⁶. Two litres of 0.2-μm filtered medium from L-Wnt3a cells, conditioned for 4 days, was adjusted to 1% Triton X-100, filtered and applied to blue (Cibacron blue) Sepharose HP (Amersham Biosciences) column (bed volume of 120 ml), which was previously equilibrated in binding buffer (150 mM KCl, 20 mM Tris-HCl, 1% CHAPS, pH 7.5). The column was then washed with four column volumes of binding buffer. Bound proteins were eluted with a single step to 1.5 M KCl, 20 mM Tris-HCl, 1% CHAPS, pH 7.5. Wnt3a eluted in two pools, each of which contained similar amounts of Wnt3a protein; however, the second pool contained significantly less total protein than the first (30.6 mg total protein in the first pool and 2.16 mg in the second pool). Fractions from this second pool were combined, concentrated to 12.5 ml on a Centricon 30 ultrafiltration device (Amicon), and fractionated on a HiLoad 26/60 Superdex 200 column (Amersham Biosciences) in phosphate buffered saline (PBS), 1% CHAPS, pH 7.3. Fractions containing Wnt3a were then fractionated on a 1-ml HiTrap Heparin column (Amersham Biosciences) in a single step elution from PBS, 1% CHAPS to PBS, 1% CHAPS, 1 M NaCl. N-terminal sequence of 1 μg purified Wnt3a was obtained by

automated Edman degradation on a Procise 494 ABI sequenator. Isolated Wnt3a begins with residue 19 of the predicted amino acid sequence (SYPIWWSLAVGPQYS), indicating that the protein is proteolytically processed to remove the signal sequence. For a detailed protocol, see <http://www.stanford.edu/~rmusse/wntwindow.html>.

Triton X-114 phase separation

Wnt3a-conditioned medium was mixed 1:1 with ice cold 4.5% Triton X-114, 150 mM NaCl, 10 mM Tris-HCl, pH 7.5, incubated on ice for 5 min, then at 31 °C for 5 min, and centrifuged at 2,000g at 31 °C for 5 min. The top, aqueous phase was separated from the bottom Triton X-114 phase and equal volumes were immunoblotted with the anti-Wnt3a antibody.

In vivo labelling of Wnt3a with palmitate

L and L-Wnt3a cells were cultured in 10-cm plates for 3 days after a 1:10 split. [9,10(n)-³H] palmitic acid (Amersham Biosciences) was added to the medium at a concentration of 0.1 mCi ml⁻¹ and incubated for 5 h at 37 °C. The medium was filtered, CHAPS was added to a concentration of 1%, and then re-filtered. The individual medium was fractionated on 1-ml HiTrap blue Sepharose columns (Amersham Biosciences) as described above. The Wnt3a-containing fractions or analogous fractions were precipitated with trichloroacetic acid and analysed by gel-electrophoresis and autoradiography.

Liquid chromatography tandem mass spectrometry

Purified Wnt3a and *Drosophila* Wnt8 were precipitated with trichloroacetic acid, re-suspended, alkylated and reduced as described²⁷. The sample was split into three aliquots, digested separately with trypsin, subtilisin and elastase, and the resulting peptide mixtures

were recombined and analysed by MudPIT as described²⁸ with modifications as described²⁷ on a Finnigan LCQ-Deca. Tandem mass spectra were searched against a database of predicted open reading frames to which common contaminants such as keratin and trypsin were added. Search results were filtered and grouped using the DTASelect program²⁹ and identifications confirmed through manual evaluation of spectra. The data were subsequently searched with a differential modification on cysteine of 238 to identify sites of palmitoylation. We also observed this peptide in its unpalmitoylated form, and at present we cannot distinguish whether the lipid is labile and lost during the manipulation of the sample or whether there is a pool of unmodified Wnt3a present in the preparation. We found the following masses [(M + H) +]: Wnt3a peptide unmodified: 1374.51 (predicted, 1374.465); Wnt3a peptide modified: 1556.10 (predicted, 1555.465); *Drosophila* Wnt8 peptide unmodified: 1583.37 (predicted, 1583.667); *Drosophila* Wnt8 peptide modified: 1764.23 (predicted, 1764.667). Although the tandem mass spectrometry analysis of Wnt3a and *Drosophila* Wnt8 identified 85% and 90% of the primary amino acid sequences, respectively, we did not obtain evidence for additional lipid modifications on other residues (S, T, Y, K, R).

Acyl-protein thioesterase treatment of Wnt3a

A total of 100 ng Wnt3a was treated in the presence of 1 µg BSA with 1, 10, 100 or 1,000 ng APT-1 (provided by A. Gilman) in buffer (PBS, 1% CHAPS, 1 M NaCl) in a total volume of 10 µl and incubated overnight at 30 °C. The reaction products were analysed in the β-catenin stabilization assay on L cells and in the Triton X-114 phase separation assay.

HSC isolation and assays

HSCs were sorted from mouse bone marrow of Bcl2 transgenic mice using antibodies as described³⁰. Cells were sorted on expression of c-Kit, Sca-1, low levels of Thy-1.1 and low to negative levels of lineage markers (Lin) using clonecyte software and the single cell deposition unit (Becton Dickinson). See Supplementary Information.

Received 12 February; accepted 20 March 2003; doi:10.1038/nature01611.

Published online 27 April 2003.

- Cadigan, K. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
- van de Wetering, M. *et al.* The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002).
- Garcia-Castro, M. I., Marcelle, C. & Bronner-Fraser, M. Ectodermal Wnt function as a neural crest inducer. *Science* **297**, 848–851 (2002).
- Papkoff, J. & Schryver, B. Secreted int-1 protein is associated with the cell surface. *Mol. Cell Biol.* **10**, 2723–2730 (1990).
- Van Leeuwen, F., Harryman Samos, C. H. & Nusse, R. Biological activity of soluble wingless protein in cultured *Drosophila* imaginal disc cells. *Nature* **368**, 342–344 (1994).
- Burrus, L. W. & McMahon, A. P. Biochemical analysis of murine Wnt proteins reveals both shared and distinct properties. *Exp. Cell Res.* **220**, 363–373 (1995).
- Hsieh, J. C., Rattner, A., Smallwood, P. M. & Nathans, J. Biochemical characterization of Wnt-frizzled interactions using a soluble, biologically active vertebrate Wnt protein. *Proc. Natl Acad. Sci. USA* **96**, 3546–3551 (1999).
- Bradley, R. S. & Brown, A. M. A soluble form of Wnt-1 protein with mitogenic activity on mammary epithelial cells. *Mol. Cell Biol.* **15**, 4616–4622 (1995).
- Roelink, H. & Nusse, R. Expression of two members of the Wnt gene family during mouse development; restricted temporal and spatial patterns in the developing neural tube. *Genes Dev.* **5**, 381–388 (1991).
- Polakis, P. Wnt signaling and cancer. *Genes Dev.* **14**, 1837–1851 (2000).
- Shibamoto, S. *et al.* Cytoskeletal reorganization by soluble Wnt-3a protein signalling. *Genes Cells* **3**, 659–670 (1998).
- Brannon, M. & Kimelman, D. Activation of Siamois by the Wnt pathway. *Dev. Biol.* **180**, 344–347 (1996).
- McKendry, R., Hsu, S. C., Harland, R. M. & Grosschedl, R. LEF-1/TCF proteins mediate wnt-inducible transcription from the *Xenopus* nodal-related 3 promoter. *Dev. Biol.* **192**, 420–431 (1997).
- Bordier, C. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* **256**, 1604–1607 (1981).
- Duncan, J. A. & Gilman, A. G. A cytoplasmic acyl-protein thioesterase that removes palmitate from G protein alpha subunits and p21(RAS). *J. Biol. Chem.* **273**, 15830–15837 (1998).
- Mason, J. O., Kitajewski, J. & Varmus, H. E. Mutational analysis of mouse Wnt-1 identifies two temperature-sensitive alleles and attributes of Wnt-1 protein essential for transformation of a mammary cell line. *Mol. Biol. Cell* **3**, 521–533 (1992).
- Maloo, J. N., Whangbo, J., Harris, J. M., Jongeward, G. D. & Kenyon, C. A Wnt signalling pathway controls hox gene expression and neuroblast migration in *C. elegans*. *Development* **126**, 37–49 (1999).
- Couso, J. P. & Arias, A. M. Notch is required for wingless signaling in the epidermis of *Drosophila*. *Cell* **79**, 259–272 (1994).
- Reya, T. *et al.* A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* advance online publication, 27 April 2003 (doi:10.1038/nature01593).
- Kadowaki, T., Wilder, E., Klingensmith, J., Zachary, K. & Perrimon, N. The segment polarity gene porcupine encodes a putative multitransmembrane protein involved in Wingless processing. *Genes Dev.* **10**, 3116–3128 (1996).
- Rocheleau, C. E. *et al.* Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* **90**, 707–716 (1997).
- Hofmann, K. A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. *Trends Biochem. Sci.* **25**, 111–112 (2000).
- Tanaka, K., Kitagawa, Y. & Kadowaki, T. *Drosophila* segment polarity gene product porcupine stimulates the posttranslational N-glycosylation of wingless in the endoplasmic reticulum. *J. Biol. Chem.* **277**, 12816–12823 (2002).
- Noordermeer, J., Klingensmith, J., Perrimon, N. & Nusse, R. *dishevelled* and *armadillo* act in the wingless signalling pathway in *Drosophila*. *Nature* **367**, 80–83 (1994).

- Pepinsky, R. B. *et al.* Identification of a palmitic acid-modified form of human Sonic hedgehog. *J. Biol. Chem.* **273**, 14037–14045 (1998).
- Wu, C. H. & Nusse, R. Ligand receptor interactions in the WNT signaling pathway in *Drosophila*. *J. Biol. Chem.* **277**, 41762–41769 (2002).
- MacCoss, M. J. *et al.* Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc. Natl Acad. Sci. USA* **99**, 7900–7905 (2002).
- Washburn, M. P., Wolters, D. & Yates, J. R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnol.* **19**, 242–247 (2001).
- Tabb, D., McDonald, W. H. & Yates, J. R. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1**, 21–26 (2002).
- Domen, J. & Weissman, I. L. Hematopoietic stem cells need two signals to prevent apoptosis; BCL-2 can provide one of these, Kit1/c-Kit signaling the other. *J. Exp. Med.* **192**, 1707–1718 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We are grateful to C.-h. Wu for generating the S2 cells expressing *Drosophila* Wnt8; M. Silverman for help in the C57MG cell transformation assay; and S. Anderson for advice on mass spectrometry. *Xenopus* embryos were provided by J. Baker and A. Borchers. A. Gilman provided APT-1, and A. Martinez-Arias the *wg*^{S21} stock. J. Nelson and members of our laboratories provided comments on the manuscript. This work was supported by Howard Hughes Medical Institute and a NIH grant awarded to T.R. R.N. is an investigator of the Howard Hughes Medical Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.N. (rnusse@cmgm.stanford.edu).

Deficiency of the adaptor SLP-65 in pre-B-cell acute lymphoblastic leukaemia

Hassan Jumaa*, Lukas Bossaller*†, Karina Portugal*†, Bettina Storch*†, Michael Lotz*, Alexandra Flemming*, Martin Schrappe‡, Ville Postila§, Pekka Riikonen||, Jukka Pelkonen§, Charlotte M. Niemeyer¶ & Michael Reth*

* Biologie III, University of Freiburg and Max Planck Institute for Immunobiology, D-79108 Freiburg, Germany

† Pediatric Hematology and Oncology, Children's Hospital, Medical School of Hanover, D-30625 Hanover, Germany

§ Department of Clinical Microbiology, University of Kuopio, PO Box 1627, and Department of Clinical Microbiology, Kuopio University Hospital, PO Box 1777, FIN-70211 Kuopio, Finland

|| Department of Pediatrics, Kuopio University Hospital, PO Box 1777, FIN-70211 Kuopio, Finland

¶ Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Freiburg University Hospital, D-79106 Freiburg, Germany

† These authors contributed equally to this work

Acute lymphoblastic leukaemia (ALL) is the commonest form of childhood malignancy, and most cases arise from B-cell clones arrested at the pre-B-cell stage of differentiation^{1,2}. The molecular events that arrest pre-B-cell differentiation in the leukaemic pre-B cells have not been well characterized. Here we show that the differentiation regulator SLP-65 (an adaptor protein also called BLNK or BASH^{3–6}) inhibits pre-B-cell leukaemia in mice. Reconstitution of SLP-65 expression in a SLP-65^{−/−} pre-B-cell line led to enhanced differentiation *in vitro* and prevented the development of pre-B-cell leukaemia in immune-deficient mice. Tyrosine 96 of SLP-65 was required for this activity. The murine SLP-65^{−/−} pre-B-cell leukaemia resembles human childhood pre-B ALL. Indeed, 16 of the 34 childhood pre-B ALL samples that were tested showed a complete loss or drastic reduction of SLP-65 expression. This loss is probably due to the incorporation of alternative exons into SLP-65 transcripts, leading to premature

stop codons. Thus, the somatic loss of SLP-65 and the accompanying block in pre-B-cell differentiation might be one of the primary causes of childhood pre-B ALL.

Mice deficient for SLP-65 show a partial block of pre-B-cell development in the bone marrow and smaller numbers of mature B cells in the spleen^{7–10}. *In vitro*, SLP-65^{−/−} pre-B cells display enhanced proliferation, and pre-B-cell leukaemia is observed in 5–10% of SLP-65^{−/−} mice¹¹. To test the requirements for the increased growth and the tumour development of SLP-65^{−/−} pre-B cells *in vivo*, we used a SLP-65^{−/−} pre-B-cell line that had been cultured for several months. This cell line and the tumours found in the SLP-65^{−/−} mice expressed the pre-B-cell receptor (pre-BCR) and were of clonal origin because both display two prominent JH recombinations (Supplementary Fig. 1). The cells were transduced with retroviral vectors coding for the green fluorescent protein (GFP) or a GFP–SLP-65 fusion protein (Fig. 1a). In the latter vector, we introduced tyrosine-to-phenylalanine mutations at either Tyr 52 (Y52F) or Tyr 96 (Y96F), the

latter being required for the binding of Bruton's tyrosine kinase (Btk) to SLP-65 (ref. 12). Expression of the Y96F GFP–SLP-65 mutant in SLP-65^{−/−} pre-B cells failed to induce pre-BCR down-regulation and kappa recombination (differentiation indicators¹¹) and thus was similar to the construct expressing only GFP (Fig. 1b, c). However, the Y52F GFP–SLP-65 mutant was similar to GFP–SLP-65 in inducing pre-BCR down-regulation and kappa recombination (Fig. 1b, c). Thus, the interaction between Btk and SLP-65 seems to be essential for SLP-65-mediated differentiation *in vitro*. To confirm these experiments *in vivo*, we established an animal model in which SLP-65^{−/−} pre-B cells were injected intravenously into immune-deficient RAG2/γc^{−/−} mice lacking any B, T and NK cells¹³. Injection of SLP-65^{−/−} pre-B cells (5 × 10⁶ per mouse) resulted in splenomegaly and leukaemia 3–5 weeks after injection (Fig. 1d and Supplementary Table).

To prove that SLP-65 is essential to limit pre-B-cell proliferation, we compared the expansion *in vivo* of SLP-65^{−/−} pre-B cells expressing GFP, GFP–SLP-65, or the Y52F or Y96F mutant proteins (Fig. 1d, e). All pre-B cells that do not differentiate *in vitro*, namely SLP-65^{−/−} pre-B cells left untreated or reconstituted with vectors for GFP or the Y96F GFP–SLP-65 mutant, proliferated and led to splenomegaly and leukaemia in RAG2/γc^{−/−} mice. In contrast, differentiation-competent pre-B cells expressing GFP–SLP-65 or the Y52F mutant form did not lead to splenomegaly or other signs of leukaemia 4 weeks after injection, although the injected cells were still detectable (Fig. 1d, e). In comparison with GFP, the pre-B cells from mice that were injected with GFP–SLP-65 transduced cells showed a down-regulation of GFP expression. However, GFP expression in these cells was still higher than the background in untransduced cells (Fig. 1e, second and fourth panels). This down-regulation might have been due to the selection of cells with low GFP–SLP-65 expression because these cells have a higher proliferation rate.

Because it is thought that human pre-B ALL originates from B-cell precursors with a maturational arrest at the pro/pre-B-cell stage¹⁴, we asked whether pre-B ALL cells show a defect in SLP-65 expression.

We tested total bone marrow cells from childhood pre-B ALL patients for the presence of SLP-65 messenger RNA and protein. The patients had been diagnosed with CD19⁺/CD10⁺ pre-B ALL (Table 1, patients 1–10) in Freiburg, Germany. RNA was isolated from total bone marrow samples that consisted of 78–93% CD19⁺ cells and, subsequently, analysis by polymerase chain reaction with reverse transcription (RT–PCR) was performed to amplify the full-length coding sequence of SLP-65. In 6 of 10 patients, only trace amounts of SLP-65 transcripts were detected in the bone marrow samples (Fig. 2a, patients 3, 5 and 7–10). As indicated by the control amplification of the B-cell-specific *mb-1* transcript (encoding Igα), the decrease in SLP-65 transcripts is not due to an inefficient synthesis of complementary DNA (Fig. 2a, lower panel). Lysates of total bone marrow cells from patients 1–6 were analysed for SLP-65 protein expression by western blotting. Confirming the PCR data, bone marrow cells of patients 3 and 5 showed no SLP-65 protein expression (Fig. 2b). Only a faint SLP-65 signal was detected in the healthy control compared with the samples from patients 1, 2, 4 and 6 (Fig. 2b). This difference is probably due to the fact that B cells represent less than 10% of the bone marrow cells in healthy people. To compare SLP-65 expression in the B-cell populations of different patients, we performed intracellular fluorescence-activated cell sorting (FACS) analysis. This confirmed the above results in that SLP-65 expression was decreased in the leukaemic pre-B cells of sample 5 in comparison with sample 6 and the healthy control (Fig. 2b, c, upper and lower panels).

Further studies were performed with purified CD19⁺ lymphoblasts/B cells from the bone marrow and blood of nine patients diagnosed for childhood pre-B ALL in Kuopio, Finland (Table 1, patients 11–19). We used the purified cells for

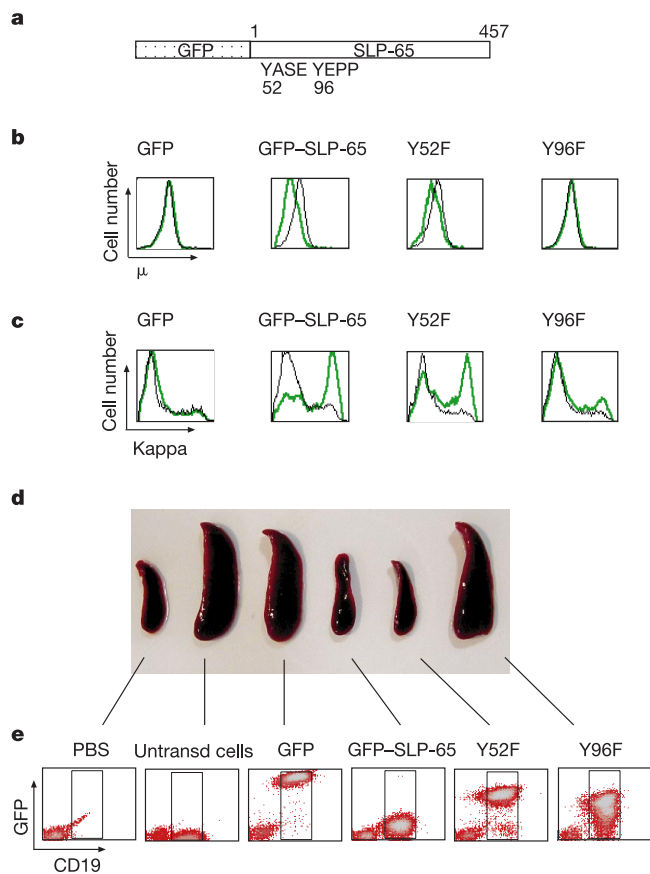


Figure 1 The Btk binding site is essential for the tumour suppressor activity of SLP-65. **a**, Structure of the GFP–SLP-65 fusion protein with tyrosine (Y) residues 52 and 96 that had been mutated to phenylalanine (F) in two different constructs. **b**, Histograms for μ expression on SLP-65^{−/−} pre-B cells 24 h after retroviral transduction of the indicated constructs (green curves) compared with untransduced cells (black curves). **c**, Histograms for kappa expression on SLP-65^{−/−} pre-B cells after retroviral transduction of the indicated constructs (green curves) compared with untransduced cells (black curves). Cells were incubated for 4 days without interleukin-7 to stimulate differentiation. **d**, Spleens of RAG2/γc^{−/−} mice 4 weeks after injection with PBS, the untransduced SLP-65^{−/−} cell line or cells transduced with GFP or the GFP–SLP-65 constructs as indicated. Before injection, the transduced cells were subjected to selection with puromycin *in vitro*. **e**, GFP–CD19 FACS profiles of the cells isolated from the corresponding spleens. Representative results of several independent experiments are shown.

RNA extraction and, in some cases, protein expression analysis. After RNA extraction and cDNA synthesis, a two-step PCR strategy was employed to achieve maximum amplification of the 5' region of the coding sequence of SLP-65. In the samples from patients 13, 16 and 17 only aberrant SLP-65 transcripts were detected at a size larger than the expected 5' region of SLP-65 (Fig. 3a). In samples 14 and 15 both aberrant and normal SLP-65 transcripts were present. Given that cells of these two samples were derived from whole blood, in which the ratio of untransformed to transformed B cells might be higher than in the bone marrow, it is conceivable that these untransformed B cells were responsible for the normal SLP-65 transcripts.

To confirm the RT-PCR results, western blot analysis was performed with cell lysates from the purified CD19⁺ cells of patients 13 and 19. As expected, no SLP-65 protein was detectable in sample 13 (Fig. 3b). Analysis of the aberrant PCR fragment amplified from patients 13–17 (Fig. 3a) revealed the presence of an additional sequence inserted between the sequences derived from exons 3 and 4 of SLP-65. By comparing the sequence of the aberrant PCR fragment with that of the human gene encoding SLP-65, we found that the additional sequence was derived from an alternative exon (which we named 3b) situated in intron 3 between exons 3 and 4. Exon 3b possesses functional 5' and 3' splice sites; however,

inclusion of this exon in the mature mRNA introduced a premature stop codon and interrupted the SLP-65 open reading frame (Fig. 3c). This could explain the low levels at which this transcript was detected because it might be degraded by nonsense-mediated decay¹⁵. Using a PCR primer pair located in exons 3 and 4, we identified another alternative exon and named it 3a (Fig. 3c). Like exon 3b, the inclusion of exon 3a interrupts the open reading frame of SLP-65. Interestingly, the murine gene encoding SLP-65 lacks sequences similar to exons 3a and 3b.

To confirm our finding that SLP-65 is missing in a substantial proportion of childhood pre-B ALL, we analysed 15 additional paediatric pre-B ALL samples by intracellular FACS for SLP-65 expression. This FACS analysis revealed that 7 of these 15 samples were deficient for SLP-65 (Table 1, patients 20–34), supporting the view that SLP-65 deficiency might be a primary cause for childhood pre-B ALL. We found a SLP-65 deficiency in one of three pre-B ALL samples derived from adult patients, whereas the remaining were positive for SLP-65 and carried a chromosomal translocation resulting in the oncogenic tyrosine kinase BCR-ABL which is thought to cause cell transformation and leukaemia¹⁶ (Table 1, patients 35–37).

The increased proliferation of the murine SLP-65^{-/-} pre-B cells requires pre-BCR expression¹¹. However, no clear correlation was observed between SLP-65 deficiency and μ -chain expression in the

Table 1 SLP-65 expression in pre-B ALL tumour samples

Patient	SLP-65 expression	Age*	B cells (%)	Cytoplasmic IgM expression	Genetic abnormality†
1	+	13.2	88	+/-	TEL/AML1
2	+	2.2	93	-	-
3	-	7.4	78	-	-
4	+	9.9	96	+	-
5	-	4.7	84	+	-
6	+	11.0	86	-	-
7	-	3.0	86	+	TEL/AML1
8	-	3.0	93	+	-
9	-	2.5	88	-	-
10	-	7.2	92	-	-
11	+	13.1	90	n.d.	Hyperdiploid‡
12	+	3.9	20	n.d.	Hyperdiploid‡
13	-	10.0	85	n.d.	-
14	+/-	0.8	52	n.d.	-
15	+/-	15.2	45	n.d.	-
16	-	8.2	85	n.d.	-
17	-	4.1	90	n.d.	TEL/AML1
18	+	7.6	90	n.d.	E2A-PBX1
19	+	11.8	90	n.d.	Hyperdiploid‡
20	+	2.9	64	+	-
21	-	3.7	89	+	-
22	+	4.0	30	+	n.d.
23	-	1.3	85	+	BCR-ABL
24	+	3.0	40	-	-
25	-	12.5	77	+	-
26	-	2.9	84	-	n.d.
27	+	10.5	91	-	-
28	-	4.8	79	-	-
29	-	5.0	75	-	-
30	+	5.0	87	-	n.d.
31	+	3.5	66	+	-
32	+	0.6	84	+	n.d.
33	-	0.2	95	-	n.d.
34	+	2.6	90	-	n.d.
35	+	43.0	68	+	BCR-ABL
36	-	29.0	92	-	-
37	+	43.0	89	+	BCR-ABL
38	+	a.b.	83	-	MLL
39	+	a.b.	96	-	MLL
40	+	10.8	98	-	MLL
41	+	17.0	97	+/-	MLL
42	+	12.0	98	+/-	MLL

SLP-65 expression in patients 1–10 and 34 (Freiburg) was analysed by FACS, RT-PCR and immunoblotting (1–6). Patients 11–19 (Kuopio, Finland) were analysed by RT-PCR and immunoblotting (13 and 19). Patients 20–33 and 41 and 42 (Hanover) were analysed by FACS and immunoblotting (41 and 42). Patients 35–37 (Frankfurt) were analysed by FACS. Patients 38–40 (Hamburg) were analysed by FACS and immunoblotting. A pre-B ALL sample was considered negative when RT-PCR and/or immunoblot analyses failed to detect the expression of functional SLP-65 (see Figs 1 and 2). For FACS analyses, a pre-B ALL sample was considered negative when the intracellular staining for SLP-65 was lower than in the B cells of a healthy control. a.b., at birth; n.d., not documented.

* Age (years) at date of diagnosis.

† In most cases, cells were analysed for the translocations BCR-ABL, TEL/AML1 and MLL rearrangements by PCR. Cytogenetic analysis was performed only for samples 11–19 and 35–39. A dash indicates that no abnormalities were detected.

‡ Possessing 55–59 chromosomes.

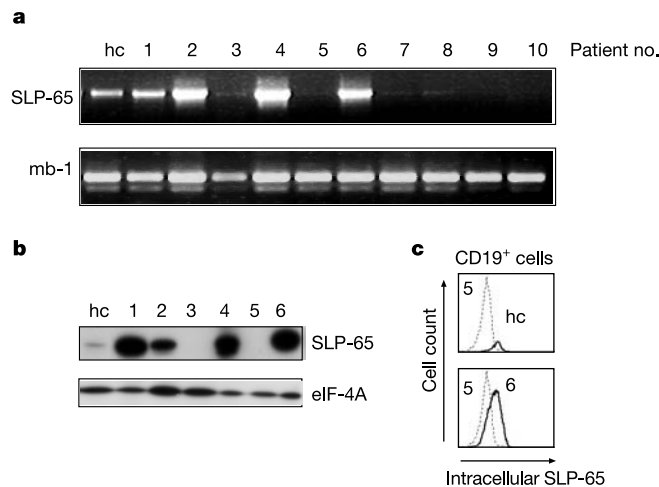


Figure 2 Absence of SLP-65 expression in samples derived from pre-B ALL patients. **a**, RT-PCR analysis of cDNA from total bone marrow cells from a healthy control (hc) and 10 patients (1–10). Primers located around the start and stop codons were used to amplify full-length SLP-65 transcripts. Transcripts from the B-cell-specific *mb-1* gene (419 and 300 bp alternative splice fragments) were used as a control for the cDNA synthesis. **b**, Western blot analysis of total bone marrow lysates from a healthy control (hc) and six patients. SLP-65 was detected with a monoclonal antibody; the detection of eIF-4A (eukaryotic initiation factor 4A) provided a loading control. **c**, FACS analysis for intracellular SLP-65 expression in B cells from patients 5 and 6 and the healthy control (hc). Only CD19-positive cells are shown.

pre-B ALL cases analysed (Table 1). A possible explanation for this observation is that an established human pre-B ALL no longer requires pre-BCR expression for its growth. The continuing RAG1/2 activity^{17,18} in the pre-B ALL cells can easily result in the loss of productive V_HDJ_H alleles, which might have been important at

the initial phase of leukaemia development. Supporting this, most pre-B ALL cells bear at least one V_HDJ_H recombination that could have undergone a V_H replacement¹⁹. The above hypothesis (that a SLP-65 deficiency combined with a pre-BCR might initiate pre-B ALL) indicates that SLP-65 deficiency might not be detected in cells transformed before V_HDJ_H recombination at the pro-B-cell stage because of, for instance, a t(4;11) translocation affecting the mixed-lineage leukaemia (MLL) gene^{20–23}. Indeed, the five analysed pre-B ALL samples with a t(4;11) translocation all expressed SLP-65 (Supplementary Fig. 2).

Our results indicate that the loss of SLP-65 expression could be due to alternative splicing of the primary SLP-65 transcript. The 12-kilobase intron between exons 3 and 4 of the human gene encoding SLP-65 contains at least two alternative exons (3a and 3b). The expression of normal SLP-65 mRNA requires skipping of these alternative exons, a process that is likely to be controlled by specific splicing factors. A disturbance of this regulation could lead to the incorporation of the alternative exons into SLP-65 mRNA, resulting in premature stop codons and the loss of SLP-65 expression from both alleles. One event that could interfere with the growth and splicing machinery of B cells is a viral infection²⁴. Although no virus has consistently been found in pre-B ALL, the hit-and-run behaviour of many viral infections could mask such associations.

Normal pre-B-cell development is characterized by a coordinated programme of differentiation, proliferation and VDJ recombination mediated by the RAG1/2 endonucleases²⁵. SLP-65^{−/−} murine pre-B cells are blocked in differentiation but show an increased proliferation and maintain RAG1/2 expression¹¹. In a recent study, the endonuclease activity of RAG2 was found to be essential for chromosomal translocations and tumour development²⁶. Thus, the continuous RAG1/2 activity found in pre-B ALL¹⁸ might explain the accumulating deletions and chromosomal translocations that further increase the diversity of pre-B ALL and make it an exceedingly heterogeneous disease.

Our finding that the loss of the adaptor protein SLP-65 is one of the primary causes of pre-B ALL might lead to a better understanding of the molecular basis of this disease. □

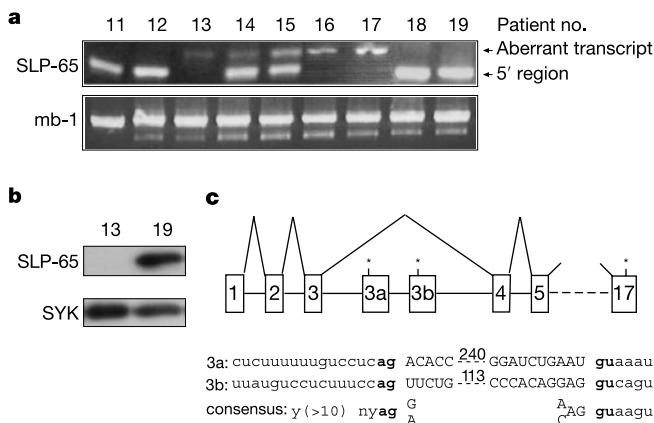


Figure 3 SLP-65 expression in CD19-purified lymphoblasts from patients with pre-B ALL. **a**, RT-PCR analysis for the expression of the 5' region of SLP-65 transcripts (exons 1–7). The B-cell-specific transcript *mb-1* was used as a control. **b**, Western blot analysis of lysates from CD19-sorted lymphoblasts of two patients. The detection of Syk provided a loading control. **c**, Exon-intron organization of the gene encoding SLP-65 (not drawn to scale). Exons are depicted as boxes, introns as lines. The incorporation of one of the alternative exons (3a or 3b) between exons 3 and 4 leads to the introduction of a premature stop codon into the SLP-65 open reading frame (indicated by asterisks). The sequences around the alternative exons 3a and 3b provide a good match to the consensus splice sequence. Exon and intron sequences are written in capital and lower-case letters, respectively. The numbers indicate the size of the corresponding exon in bp. Y, pyrimidine; N, any nucleotide; Y(n), polypyrimidine tract.

Methods

Patients

The study protocol included children diagnosed with B-lineage acute lymphoblastic leukaemia (pre-B ALL) at hospitals in Germany (Freiburg, Hanover, Hamburg) and Finland (Kuopio). The specimens from Hanover were derived from the ALL-BFM study group on the treatment of childhood ALL. The design, conduct, analysis and results of these multi-centre trials have been described in detail elsewhere²⁷.

RT-PCR analysis

RNA was purified with a MasterPure RNA purification kit (EpiCentre; Finnish patients) or Trizol (Gibco; German patients). cDNA was generated with Moloney-murine-leukaemia virus reverse transcriptase and poly(dT) primers. The following primers were used for the amplification of SLP-65 transcripts: 5'-primer, 5'-TGGACAGTTATTCGTCTCTCTT-3'; 3'-primer, 5'-CTTTGGGAAAGTGTCTGAAGC-3'; exon-7-primer, 5'-GTGAACCTGCTTCTGTGGGA-3'. For the amplification of full-length SLP-65 the 5' and 3' primers were used. For the amplification of the 5' region of SLP-65, the 5' and exon 7 primer pair was used in two rounds of amplification.

The following primers were used for the amplification of *mb-1*: 5'-primer, 5'-CTCTGCCTGCCACCATCTTC-3'; 3'-primer, 5'-CCCCTCGGCTGTGATGATTC-3'.

Western blot analysis

Lysates of total bone marrow were prepared by the Trizol method in accordance with the manufacturer's instructions. Lysates of sorted lymphoblasts were prepared by lysing 10^6 cells in lysis buffer, pH 7.4 (137.5 mM NaCl, 50 mM Tris-HCl, 0.5 mM EDTA, 1% Nonidet P40, 10% glycerol). Samples (from $(1-2) \times 10^5$ cell equivalents) were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; 10% or 12% polyacrylamide) and transferred to poly(vinylidene difluoride) membranes (Immobilon-P; Millipore) as described previously¹¹. SLP-65 was detected with the monoclonal antibody B-211 (Santa Cruz). Antibodies against eIF-4A have been described elsewhere²⁸. Immunoreactive proteins were detected by the ECL chemiluminescence detection system (Amersham).

Intracellular FACS staining

Cells were stained for surface CD19 (Pharmingen), fixed by resuspension in 4% paraformaldehyde, permeabilized by the addition of 0.1% saponin and subsequently incubated with the anti-SLP-65 antibody B-211. A fluorescein isothiocyanate-labelled goat anti-mouse IgG2a was used as a secondary antibody. FACS analysis was performed with a FACSCalibur (Becton Dickinson).

Received 30 January; accepted 28 March 2003; doi:10.1038/nature01608.

- Jennings, C. D. & Foon, K. A. Recent advances in flow cytometry: application to the diagnosis of hematologic malignancy. *Blood* **90**, 2863–2892 (1997).
- Ma, S. K., Wan, T. S. K. & Chan, L. C. Cytogenetics and molecular genetics of childhood leukemia. *Hematol. Oncol.* **17**, 91–105 (1999).
- Wienands, J. et al. SLP-65: a new signaling component in B lymphocytes which requires expression of the antigen receptor for phosphorylation. *J. Exp. Med.* **188**, 791–795 (1998).
- Fu, C., Turck, C. W., Kurosaki, T. & Chan, A. C. Blnk—a central linker protein in B cell activation. *Immunity* **9**, 93–103 (1998).
- Goitsuka, R. et al. BASH, a novel signaling molecule preferentially expressed in B cells of the bursa of Fabricius. *J. Immunol.* **161**, 5804–5808 (1998).
- Mingishi, Y. et al. An essential role for BLNK in human B cell development. *Science* **286**, 1954–1957 (1999).
- Jumaa, H. et al. Abnormal development and function of B lymphocytes in mice deficient for the signaling adaptor protein SLP-65. *Immunity* **11**, 547–554 (1999).
- Pappu, R. et al. Requirement for B cell linker protein (BLNK) in B cell development. *Science* **286**, 1949–1954 (1999).
- Hayashi, K. et al. The B cell-restricted adaptor BASH is required for normal development and antigen receptor-mediated activation of B cells. *Proc. Natl Acad. Sci. USA* **97**, 2755–2760 (2000).
- Xu, S. L. et al. B cell development and activation defects resulting in xid-like immunodeficiency in BLNK/SLP-65-deficient mice. *Int. Immunol.* **12**, 397–404 (2000).
- Flemming, A., Brummer, T., Reth, M. & Jumaa, H. The adaptor protein SLP-65 acts as a tumour suppressor that limits pre-B cell expansion. *Nature Immunol.* **4**, 38–43 (2003).
- Chiu, C. W., Dalton, M., Ishiai, M., Kurosaki, T. & Chan, A. C. BLNK: molecular scaffolding through 'cis'-mediated organization of signaling proteins. *EMBO J.* **21**, 6461–6472 (2002).
- Disanto, J. P., Muller, W., Guygrand, D., Fischer, A. & Rajewsky, K. Lymphoid development in mice with a targeted deletion of the interleukin 2 receptor gamma chain. *Proc. Natl Acad. Sci. USA* **92**, 377–381 (1995).
- Goodman, P. A., Wood, C. M., Vassilev, A., Mao, C. & Uckun, F. M. Spleen tyrosine kinase (Syk) deficiency in childhood pro-B acute lymphoblastic leukemia. *Oncogene* **20**, 3969–3978 (2001).
- Hillgren, P. & Parker, R. Mechanisms of mRNA surveillance in eukaryotes. *Annu. Rev. Genet.* **33**, 229–260 (1999).
- Advani, A. S. & Pendergast, A. M. Bcr–Abl variants: biological and clinical aspects. *Leuk. Res.* **26**, 713–720 (2002).
- Umiel, T., Pattengale, P. & Weinberg, K. Recombination activating gene-1 (RAG-1) expression in all differentiation stages of B-lineage precursor acute lymphoblastic leukemia. *Leukemia* **7**, 435–440 (1993).
- Nishii, K. et al. Expression of B cell-associated transcription factors in B-cell precursor acute lymphoblastic leukemia cells: association with PU.1 expression, phenotype, and immunogenotype. *Int. J. Hematol.* **71**, 372–378 (2000).
- Beishuizen, A. et al. Multiple rearranged immunoglobulin genes in childhood acute lymphoblastic leukemia of precursor B-cell origin. *Leukemia* **5**, 657–667 (1991).
- Gu, Y. et al. The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila* trithorax, to the AF-4 gene. *Cell* **71**, 701–708 (1992).
- Tkachuk, D. C., Kohler, S. & Cleary, M. L. Involvement of a homolog of *Drosophila* trithorax by 11q23 chromosomal translocations in acute leukemias. *Cell* **71**, 691–700 (1992).
- Domer, P. H. et al. Acute mixed-lineage leukemia t(4;11)(q21;q23) generates an MLL–AF4 fusion product. *Proc. Natl Acad. Sci. USA* **90**, 7884–7888 (1993).
- Armstrong, S. A. et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genet.* **30**, 41–47 (2002).
- Nordqvist, K., Ohman, K. & Akusjarvi, G. Human adenovirus encodes two proteins which have opposite effects on accumulation of alternatively spliced mRNAs. *Mol. Cell. Biol.* **14**, 437–445 (1994).
- Rajewsky, K. Clonal selection and learning in the antibody system. *Nature* **381**, 751–758 (1996).
- Zhu, C. M. et al. Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* **109**, 811–821 (2002).
- Schrappe, M. et al. Long-term results of four consecutive trials in childhood ALL performed by the ALL-BFM study group from 1981 to 1995. *Leukemia* **14**, 2205–2222 (2000).
- Ederly, I. et al. Involvement of eukaryotic initiation factor 4A in the cap recognition process. *J. Biol. Chem.* **258**, 11398–11403 (1983).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank E. Bergsträsser (Zürich), U. zur Stadt (Hamburg), J. Harbott (Giessen) and O. G. Ottmann (Frankfurt) for pre-B ALL samples; W.-D. Ludwig (Berlin) and J. Harbott (Giessen) for immunophenotyping and molecular genetic studies of the patients from Freiburg and Hanover; P. Nielsen for reading the manuscript and H. Mossmann for help in the mouse experiments; and C. Eschbach and U. Stauffer for technical support. Financial support for these experiments was provided by the Deutsche Forschungsgemeinschaft and the Leibniz programme to M.R.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.R. (reth@immunbio.mpg.de) or H.J. (jumaa@immunbio.mpg.de).

Humanin peptide suppresses apoptosis by interfering with Bax activation

Bin Guo, Dayong Zhai, Edelmira Cabezas, Kate Welsh, Shahrzad Nouraini, Arnold C. Satterthwait & John C. Reed

The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, California 92037, USA

Bax (Bcl2-associated X protein) is an apoptosis-inducing protein that participates in cell death during normal development and in various diseases¹. Bax resides in an inactive state in the cytosol of many cells. In response to death stimuli, Bax protein undergoes conformational changes that expose membrane-targeting domains, resulting in its translocation to mitochondrial membranes, where Bax inserts and causes release of cytochrome *c* and other apoptogenic proteins². It is unknown what controls conversion of Bax from the inactive to active conformation. Here we show that Bax interacts with humanin (HN), an anti-apoptotic peptide of 24 amino acids encoded in mammalian genomes^{3,4}. HN prevents the translocation of Bax from cytosol to mitochondria. Conversely, reducing HN expression by small interfering RNAs sensitizes cells to Bax and increases Bax translocation to membranes. HN peptides also block Bax association with isolated mitochondria, and suppress cytochrome *c* release *in vitro*. Notably, the mitochondrial genome contains an identical open reading frame, and the mitochondrial version of HN can also bind and suppress Bax. We speculate therefore that HN arose from mitochondria and transferred to the nuclear genome, providing a mechanism for protecting these organelles from Bax.

To identify proteins that might bind selectively to inactive Bax, we performed a yeast two-hybrid screen of a human testis complementary DNA library using as a bait human Bax(S184K)—a mutant form of Bax locked into an inactive conformation⁵. One cDNA clone confirmed to interact specifically with Bax(S184K) contained an in-frame cDNA encoding HN.

HN was originally identified as an anti-apoptotic peptide encoded in a cDNA that rescued neuronal cells from apoptosis induced by presenilin mutants associated with familial Alzheimer's disease and by amyloid- β protein³. Mutagenesis studies showed that the cysteine at position 8 of the HN peptide is critical for its anti-apoptotic function^{3,4}. We therefore contrasted the ability of wild-type and Cys8Pro (C8P) mutant HN to interact with Bax in co-immunoprecipitation assays, expressing HN peptides as fusions with green fluorescent protein (GFP) in human cells. GFP–HN co-immunoprecipitated Myc-tagged Bax from HEK293T cells, but not GFP or GFP–HN(C8P) (Fig. 1a). Immunoblot analysis confirmed production of all proteins. GFP–HN also co-immunoprecipitated endogenous Bax from certain cell lines (not shown).

We surveyed multiple Bcl2-family proteins for interactions with HN by co-immunoprecipitation experiments (Fig. 1a). GFP–HN did not co-immunoprecipitate with other Bcl2-family proteins that are predicted to share structural similarity with Bax, including Bcl2, Bcl-xL, Bak, Bcl-B, Mcl1 and Bok⁶.

Expression of endogenous HN has been demonstrated in the testis and colon of mice⁷. Using a rabbit polyclonal antibody raised against HN peptide (gift of I. Nishimoto) for co-immunoprecipitations, we determined that endogenous HN peptide interacts with endogenous Bax in mouse testis. Endogenous HN was also determined to be a predominantly cytosolic protein (see Supplementary Information).

To characterize further the binding of HN to Bax, we generated an *in vitro* binding assay using the fluorescence polarization technique.

Full-length Bax protein was produced in bacteria and purified⁸, and was then incubated with rhodamine-conjugated synthetic purified HN peptide. As shown in Fig. 1c, Bax bound rhodamine–HN in a concentration-dependent and saturable manner, with an estimated dissociation constant (K_d) of 2 nM. In contrast, various control peptides of similar length, such as rhodamine–CD40 (residue P250 to G266), did not display interactions with Bax in fluorescence polarization assays.

We next performed various cell-based experiments to explore the functional significance of HN-binding to Bax (Fig. 2; see also Supplementary Information). First, we tested the effects of HN in yeast (*Saccharomyces cerevisiae*), where ectopic expression of Bax induces cell death through a mechanism similar to mammalian cells (reviewed in ref. 9). We expressed Bax using a conditional promoter, where Bax is produced when yeast are plated on galactose- but not glucose-containing medium¹⁰. Co-expression of wild-type HN peptide, expressed as a fusion protein with TAD, rescued yeast from Bax-induced lethality, whereas HN(C8P) mutant did not (Fig. 2a). Immunoblot analysis demonstrated that both the HN and

HN(C8P) fusion proteins were produced at comparable levels in yeast, excluding differences in expression as a trivial explanation for the findings (not shown).

Next, we examined the effects of HN overexpression in mammalian cells using several cell death stimuli known to induce apoptosis partly through Bax-dependent mechanisms¹¹. When Flag-tagged HN or Flag-tagged control (27-amino-acid) peptides were expressed in CSM14.1 cells (immortalized rat hippocampal neurons)¹², apoptosis induced by kinase inhibitor (staurosporine (STS)), serum deprivation and ultraviolet irradiation was selectively suppressed by HN–Flag (Fig. 2b, c). Conversely, apoptosis induced by a Bax-independent death stimulus, tumour necrosis factor (TNF), was not suppressed (Fig. 2c). Although HN reduced by half apoptosis induced by 0.2 μ M STS ($52 \pm 4\%$ compared with $28 \pm 3\%$; $P = 0.03$), HN-mediated protection was overcome at higher doses of STS (1 μ M) (Fig. 2b), consistent with invoking Bax-independent (HN-insensitive) mechanisms of apoptosis at higher doses¹¹.

As an alternative to over-expression studies, synthetic small interfering RNA (siRNA)¹³ was used to knock down expression of endogenous HN in SF268 cells, a glioblastoma cell line empirically determined to contain high levels of endogenous HN. Transfection into SF268 cells of HN siRNA, but not control siRNAs, reduced endogenous levels of HN, as determined by immunoblotting, correlating with increased sensitivity to apoptosis induced by STS and serum-deprivation (mitochondrial-dependent) but not by TNF-related apoptosis-inducing ligand (TRAIL; mitochondria-independent) (Fig. 2d). The specificity of the siRNA-mediated effect was further confirmed by various control transfections, which did not nonspecifically sensitize cells to apoptosis (Supplementary Information). Again, whereas HN siRNA doubled the percentage of cell death in cultures exposed to low doses of STS ($53 \pm 4\%$ compared with $27 \pm 3\%$), this sensitization was overcome by high doses of STS (Fig. 2d), consistent with the studies of HN overexpression discussed above.

As HN immunoprecipitates with Bax but not Bak, we next contrasted the effects of HN on apoptosis induced by these two Bcl2-family proteins; performing transient transfection experiments using CSM14.1 neurons or prostate cancer PC-3 cells. HN suppressed by half the percentage of apoptosis induced by Bax, whereas Bak-induced apoptosis was unaffected (Fig. 2e).

For the next experiment investigating the functional significance of HN-binding to Bax, we contrasted the effects of HN in Bax-expressing and Bax-deficient human cells, reasoning that if the mechanism by which HN suppresses apoptosis involves Bax binding, then HN should not protect Bax-deficient cells. For these experiments, we used HCT116 colon cancer cells, which contain one intact and one mutant *BAX* allele, and a mutant of HCT116 (gift of B. Vogelstein) in which the remaining allele was disrupted by homologous recombination, producing Bax-deficient cells¹⁴. In contrast to their differences in Bax expression, these cell lines both express comparable amounts of Bid (Fig. 2f) and several other Bcl2-family proteins (not shown). In HCT116 parental cells, HN–Flag significantly reduced the percentage of cells undergoing STS-induced apoptosis, suppressing apoptosis by approximately half when modest doses of STS were used ($62 \pm 5\%$ compared with $32 \pm 3\%$; $P = 0.002$) (Fig. 2f). In contrast, HN–Flag failed to suppress apoptosis in Bax-deficient HCT116 cells. On the basis of previous studies using cells from gene knockout mice that have shown that either Bax or Bak is sufficient for STS-induced apoptosis¹¹, we presume the reason why HN-mediated protection is overcome at high doses of STS in parental HCT116 cells is because HN does not interfere with Bak.

Finally, using a series of HN mutants, we observed a perfect correlation between Bax binding and suppression of apoptosis induced by overexpression of Bax (Supplementary Information).

To explore the mechanism by which HN suppresses apoptosis induced by Bax, we examined the effects of HN overexpression on

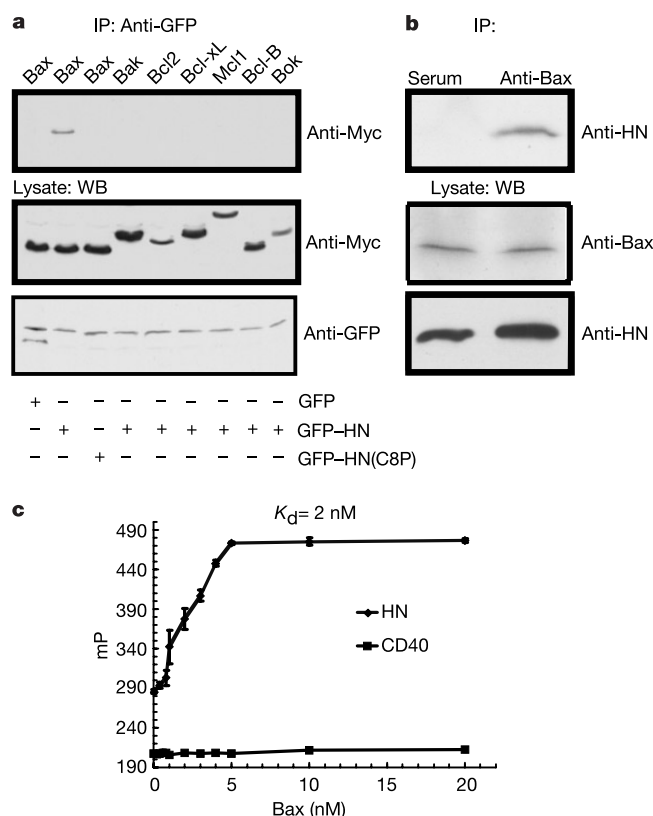


Figure 1 HN interacts with Bax. **a**, HEK293T cells were transfected with pcDNA3-Myc plasmids encoding Bax or other Bcl2-family proteins with GFP, GFP–HN or GFP–HN(C8P). Cell lysates were immunoprecipitated with polyclonal anti-GFP antibody. The immunoprecipitates (IP) or the lysates were blotted with anti-Myc or anti-GFP antibodies, respectively. The slower migrating form of GFP seen in lane 1 represents an alternative form of GFP produced from the wild-type (control) p-EGFP-C1 plasmid (<http://www.clontech.com>). **b**, Testes of three-week-old mice were homogenized in lysis buffer (0.1% Triton X-100, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, plus protease inhibitor cocktail). The lysates were immunoprecipitated with control rabbit serum or polyclonal anti-mouse Bax antibody. The immunoprecipitates or the lysates were subjected to electrophoresis in Tris/Tricine gels (16%T, 3%C), electroblotted to PVDF membrane, and immunoblotted with polyclonal anti-HN antibody or monoclonal anti-Bax antibody 6A7 (Trevigen). **c**, The affinity (K_d) of Bax–HN interaction was measured by fluorescence polarization assay, using various concentrations of purified recombinant Bax and 40 nM rhodamine-conjugated HN peptide. Control peptide from CD40 is also shown (mean $\pm$ s.d.; $n = 3$).

STS-induced translocation of GFP-Bax protein from cytosol to mitochondria in cells^{2,5,15}. For these experiments, synthetic HN or HN(C8P) peptides were introduced into cells¹⁶, and we monitored the translocation of GFP-Bax to mitochondria by microscopy^{2,5,15}, determining the percentage of cells in which cytosolic fluorescence was diffuse versus punctate. Treatment of GFP-Bax-expressing cells

with STS induced mitochondrial translocation of GFP-Bax in most cells (Fig. 3a; see also Supplementary Information), whereas GFP-Bax translocation was suppressed by about half in cells transduced with HN but not HN(C8P) control peptide. These findings were also confirmed by subcellular fractionation, where cytosol and mitochondria-enriched heavy membrane fractions were

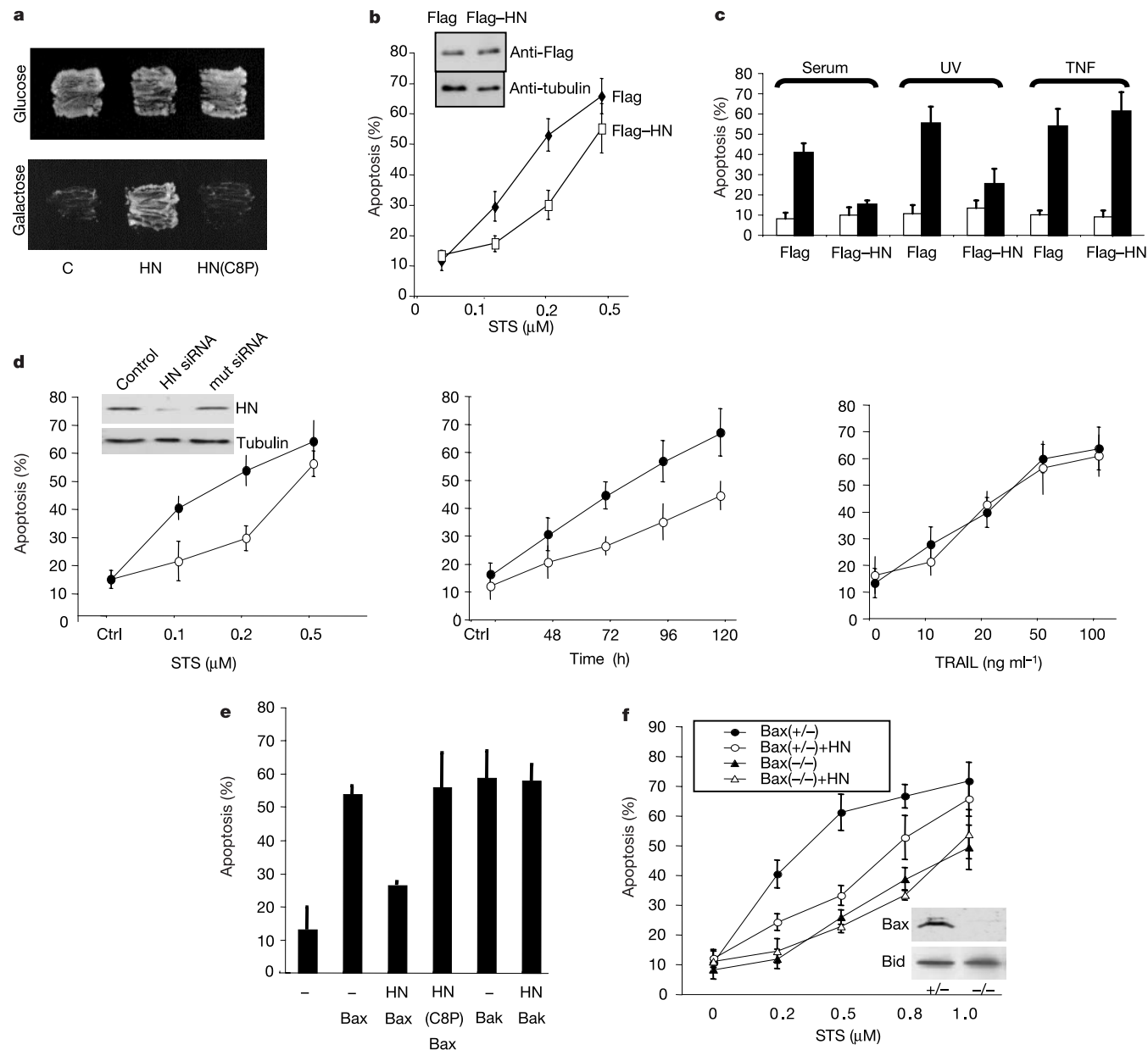


Figure 2 HN inhibits cell death induced by Bax. **a**, HN prevents Bax-induced cell death in yeast. Cells were co-transformed with YEp51-Bax (encoding Bax under control of a *GAL10* promoter) and plasmid control (C) or plasmids encoding HN or HN(C8P). Cells grown in glucose-containing medium were streaked onto either glucose or galactose plates to suppress or induce Bax expression, respectively. **b**, CSM14.1 neuronal cells were transfected with plasmids encoding Flag-HN or Flag-control peptide, together with GFP-encoding plasmid at a 4:1 ratio. Lysates prepared from replicate cultures of transfected cells were normalized for total protein content and analysed by immunoblotting using anti-Flag and anti-tubulin antibodies, confirming production of similar amounts of Flag-control and Flag-HN peptides (inset). **c**, CSM14.1 cells transfected as above were cultured either without further treatment (open bars) or subjected to apoptotic stimuli (filled bars), including serum deprivation for 3 days, 10 J m⁻² ultraviolet irradiation, followed by 24 h culture, or 50 ng ml⁻¹ TNF plus 1 μ M 2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid (CDDO) for 24 h. Apoptosis (%) was

determined by DAPI staining. **d**, SF268 cells were transfected with HN siRNAs (filled circles) or scrambled control siRNAs (open circles). After 72 h, cell lysates were prepared and 100 μ g was analysed by immunoblotting using anti-HN and anti-tubulin (as a loading control) antibodies (inset). (See Methods for full details.) **e**, HN protects against Bax- but not Bak-induced apoptosis. CSM14.1 cells were co-transfected with pcDNA3-HA-Bax or pcDNA3-HA-Bak together with plasmids encoding with GFP (-), GFP-HN or GFP-HN(C8P). Apoptosis was measured 48 h later among GFP⁺ cells. **f**, HN inhibits STS-induced apoptosis in wild-type HCT116 cells, but not HCT116 BAX^{-/-} cells. HCT116 cells heterozygous or homozygous for BAX gene inactivation were transfected with plasmids encoding Flag (filled symbols) or HN-Flag (open symbols), together with GFP-encoding plasmids at a 4:1 ratio. After 24 h, cells were cultured with or without 0.2–1.0 μ M STS for 8 h, then apoptosis was determined for GFP⁺ cells (mean $\pm$ s.d.; $n = 3$). Lysates from the cells were analysed by immunoblotting (inset; 20 μ g) using anti-Bax and anti-Bid antibodies.

prepared from cells transduced with untagged HN or HN(C8P) peptides. We measured the relative amounts of endogenous Bax protein in these two fractions by immunoblotting (Fig. 3b).

The peptide transfer findings were extended to address the role of endogenous HN using siRNA methods. In SF268 cells, which contain high levels of endogenous HN peptide (unlike Cos7 cells), most of the Bax protein remained in the cytosol after treatment with low doses of STS (0.1 μ M) (Fig. 3c). In contrast, when HN expression was knocked down by siRNA in SF268 cells, then STS-induced translocation of Bax to membranes was enhanced (Fig. 3c). We conclude therefore that HN suppresses translocation of Bax to mitochondria.

It has been suggested that HN might block apoptosis by acting outside cells, through HN-binding cell surface receptors, rather than inside cells^{3,4}. To determine whether HN can act directly on Bax, we tested the effects of HN on isolated mitochondria. Addition of recombinant purified Bax protein to mitochondria *in vitro* induces cytochrome *c* release¹⁷. Pre-incubating Bax with HN peptide suppressed Bax association with mitochondria and reduced cytochrome *c* release (Fig. 3d). In contrast, the non-Bax-binding HN(C8P) mutant peptide did not interfere with the effects of Bax on isolated mitochondria. We conclude that HN can directly suppress the targeting of Bax to mitochondria. Furthermore, *in vitro* protein binding experiments suggested that HN stabilizes the conformation of Bax in which the carboxy-terminal hydrophobic α -helical transmembrane domain is docked onto the body of the Bax protein (Supplementary Information).

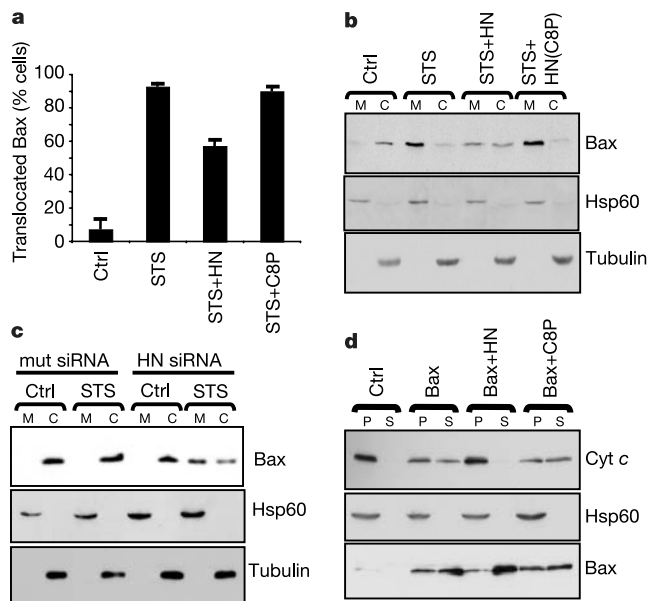


Figure 3 HN blocks Bax translocation to mitochondria. **a**, HN peptide prevents STS-induced Bax translocation *in vivo*. Cos7 cells were first transfected with GFP–Bax plasmids. After 24 h, untagged wild-type or HN(C8P) mutant peptides were introduced into cells using Chariot reagent, and 2 h later cells were treated with 1 μ M STS. After 4 h, the percentage of cells with translocated Bax was determined (mean $\pm$ s.d.; $n = 3$) by confocal microscopy. **b**, Wild-type or HN(C8P) mutant peptides were introduced into Cos7 cells using Chariot reagent. Cytosolic (C) and mitochondria-containing (M) fractions were isolated by differential centrifugation²² and analysed by immunoblotting using anti-Bax (top), Hsp60 (middle; mitochondrial marker), or tubulin (bottom; cytosol marker) antibodies. **c**, HN or mutant control double-stranded siRNAs were transfected into SF268 cells. After 48 h, cells were cultured without (control) or with 0.1 μ M STS. After 24 h, cells were fractionated and analysed by immunoblotting as above. **d**, Isolated mitochondria were mixed with 400 ng Bax protein, with or without pre-incubating Bax with 100 μ M HN or HN(C8P) peptides for 10 min. After 3 h at 30 $^{\circ}$ C, mitochondria were centrifuged and the resulting pellets (P) and supernatants (S) were analysed by immunoblotting using anti-cytochrome *c* (top), Hsp60 (middle) and Bax (bottom) antibodies.

During analysis of HN-encoding sequences in the human genome, we noticed an identical open reading frame (ORF) embedded in the 16S ribosomal RNA gene of the mammalian mitochondrial genome (Fig. 4a). Differences in codon usage by the endogenous protein translation machinery of mitochondria predict a slightly different HN peptide (Fig. 4b), if produced within these organelles¹⁸. We therefore contrasted the ability of the predicted nuclear-encoded and mitochondria-encoded HN peptides (termed HN(N) and HN(M), respectively) to bind Bax and to suppress apoptosis induced by overexpression of Bax. Using HN(N) and HN(M) fused to GFP, comparable amounts of Bax co-immunoprecipitated with both proteins (Fig. 4c). Bax-induced apoptosis was also suppressed to comparable extents by GFP–HN(N) and GFP–HN(M) (Fig. 4d). Thus, the nuclear and the mitochondrial translations of the HN ORF bind and suppress Bax.

It was proposed previously that the HN peptide is secreted from cells, acting on neighbouring cells³. We demonstrate here that Bax is the molecular target of HN relevant to its anti-apoptotic phenotype. Mutants of HN that fail to bind Bax also fail to suppress apoptosis. Moreover, HN peptide blocks the effects of Bax on isolated

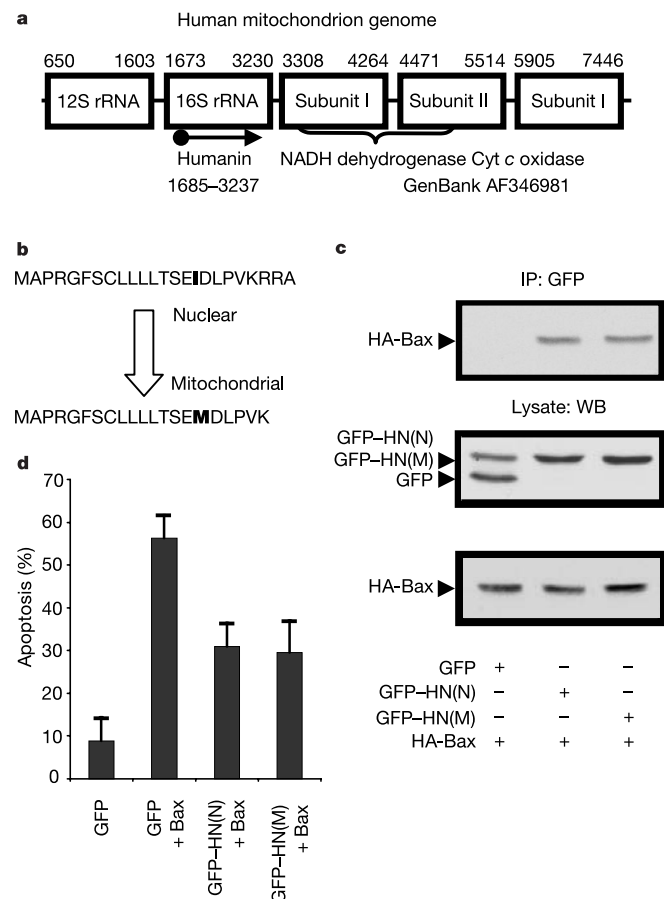


Figure 4 Comparison of nuclear (N)- and mitochondria (M)-encoded HN. **a**, An ORF identical in nucleotide sequence to the previously reported HN cDNA³ is embedded in the 16S rRNA gene in the mitochondrial genome. **b**, The amino-acid sequence of the putative mitochondrial HN peptide was predicted according to mitochondrial codon usage. **c**, HEK293T cells were transfected with pcDNA3–HA–Bax together with plasmids encoding GFP, GFP–HN(N) or GFP–HN(M). Cell lysates were immunoprecipitated with polyclonal anti-GFP antibody. Immunoprecipitates (IP) or cell lysates were analysed by immunoblotting using anti-haemagglutinin (HA) or anti-GFP antibodies, respectively. **d**, CSM14.1 neuronal cells were co-transfected with pcDNA3–HA–Bax together with plasmids encoding GFP, GFP–HN(N) or GFP–HN(M). The percentage of apoptosis was determined 48 h later by DAPI staining (mean $\pm$ s.d.; $n = 3$).

mitochondria, excluding a need for a plasma membrane receptor (also the case in yeast), thus dissociating its mechanism of action from other mammalian proteins.

The HN peptide is encoded in several long messenger RNAs that have been detected *in vivo*, and the endogenous HN peptide, with a relative molecular mass of 3,000, has been detected in certain normal tissues⁷. Using a BLAST search, we found cDNAs identical or similar to HN in plants, nematodes, rats, mice and many other species (see Supplementary Information). Although we have no reason to suspect this ORF is expressed in mitochondria currently, the presence of an identical HN-encoding ORF embedded in the mitochondrial genome raises the possibility that replicas of this ORF were transferred during evolution from these organelles into mammalian genomes. Given that mitochondria are thought to have arisen from bacteria that established residence in the cytosol of eukaryotic cells¹⁹, it is also intriguing that the Bax protein shares marked structural similarity with the pore-forming domains of bacterial colicin proteins, molecules which are secreted and used as weapons to kill competing bacteria (reviewed in ref. 20). Although the particular mechanisms differ, the ability of HN peptides to suppress Bax translocation to mitochondrial membranes is reminiscent of the immunity proteins (such as Cai and Cbi) of bacteria that bind and inhibit colicins²¹. HN, however, shares no sequence homology with these bacterial proteins.

Although additional mechanisms besides HN binding may be involved in the control of Bax translocation *in vivo*, determination of the structure of HN bound to Bax may suggest strategies for developing small-molecule drugs that inhibit Bax. In this regard, Bax has been implicated through gene ablation studies in mouse models in numerous diseases associated with pathological cell loss, including stroke, Parkinson's disease, and oocyte depletion during menopause, making it a promising target for new therapies. □

Methods

Yeast two-hybrid assays

Yeast two-hybrid screening of cDNA libraries was performed using hu-Bax(S184K) in pGilda. A human adult testis cDNA library (Invitrogen) was screened using standard yeast two-hybrid procedures (see Supplementary Information).

Cell culture, transfections, imaging and apoptosis assays

CSM14.1, HCT116, Cos7 and SF268 cells were cultured in DMEM high-glucose medium (Irvine Scientific) and PC-3 cells in RPMI 1640 medium, containing 10% fetal bovine serum (FBS). Transfection of cells was performed using SuperFect (Qiagen) or LipofectaminePLUS reagent (Invitrogen). CSM14.1 cells were cultured at 39 °C after transfection to inactive temperature-sensitive large T antigen, as described¹². In Fig. 2b, CSM14.1 cells were cultured for 8 h with 0.1–0.5 μM STS, then fixed and stained with 4,6-diamidino-2-phenylindole (DAPI) to determine the percentage of apoptotic cells, evaluating ≥200 GFP⁺ cells per sample (mean ± s.d.; *n* = 3). We performed confocal microscopy and apoptosis assays as described^{15,22} (see also Supplementary Information). In Fig. 2d, the percentage of apoptosis was measured after the following treatments: (1) SF268 cells cultured with various concentrations of STS for 8 h, at 72 h after siRNA transfection (mean ± s.d.; *n* = 3) (left panel); (2) SF268 cells washed 4 h after transfection with HN siRNA or control dsRNA and then cultured for 2–5 days without serum (middle panel); (3) 3 days after transfection, SF268 cells were treated with 10–100 ng ml⁻¹ TRAIL (Alexis) plus 25 μg ml⁻¹ cycloheximide for 8 h (right panel).

Immunoblotting and immunoprecipitations

Immunoblotting was performed as described previously²². For co-immunoprecipitations, cells were cultured in 50 μM benzoyloxycarbonyl-Val-Ala-Asp (O-methyl)-fluoromethyl ketone (zVAD-fmk) to prevent apoptosis. Cells were suspended in lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 20 mM EDTA, 50 mM NaF, 0.5% NP-40, 0.1 mM Na₃VO₄, 20 μg ml⁻¹ leupeptin, 20 μg ml⁻¹ aprotinin, 1 mM dithiothreitol and 1 mM phenylmethyl sulphonyl fluoride (PMSF)). Lysates (200 μl diluted in 1 ml final volume of lysis buffer) were cleared by incubation with 15 μl protein G Sepharose 4B (Zymed) and then incubated with 15 μl polyclonal antibody and 15 μl protein G at 4 °C overnight. Beads were then washed four times with 1.5 ml lysis buffer before boiling in Laemmli sample buffer and performing SDS-polyacrylamide gel electrophoresis or immunoblotting.

Fluorescence polarization assays

Recombinant Bax protein was isolated from *Escherichia coli* BL21 harbouring pTYB1-Bax essentially as described⁸. For fluorescence polarization assays, various concentrations of Bax protein were incubated with 40 nM of rhodamine-conjugated synthetic purified HN peptide dissolved in water for 30 min in dark. Fluorescence polarization was measured using an Analyst AD Assay Detection System (LJL Biosystem).

Peptides

Rhodamine-conjugated HN peptide and unconjugated HN and HN(C8P) peptides were synthesized using fluorenylmethoxycarbonyl chemistry with DIC coupling²³ (see Supplementary Information).

Peptide transfections

HN or HN(C8P) peptides were transfected into GFP-Bax-transfected Cos7 cells using Chariot reagent (Active Motif)¹⁶. One microgram of peptide was mixed with 6 μl of Chariot reagent in 200 μl water and incubated for 30 min. Two hours before STS treatment, GFP-Bax-expressing cells in 6-well plates were washed with PBS and incubated with Chariot-peptide complex in serum-free medium at 37 °C for 1 h. Cells were incubated for an additional hour after 1 ml complete growth medium was added. Cells were then treated with STS to induce Bax translocation. Pilot experiments with fluorochrome-conjugated peptide indicated >80% cell transfection.

Preparation and transfection of siRNA

Two different double-stranded ribo-oligonucleotides with overhanging 3' deoxy TT were prepared that target HN mRNAs (HN siRNA1: r(CCAGUGAAAUUGACCUGCC)d(TT); HN siRNA2: r(GGGUUCAGCUGUCUUCUAC)d(TT)), along with two scrambled controls (mut siRNA1: r(UUCGACCUAGGAUCACACG)d(TT); mut siRNA2: r(CGCAUCUGUUAUCGGUGUC)d(TT)). Double-stranded (ds)RNAs were dissolved in sterile 100 mM potassium acetate, 30 mM HEPES-KOH, 2 mM magnesium acetate, pH 7.4, to a final concentration of 20 μM. SF268 cells were cultured in 6-well plates in 2 ml DMEM medium containing 10% FBS. Cells at 40% confluency were transfected with a 1:1 mixture of both siRNAs or both control dsRNAs using 10 μl oligofectamine (Invitrogen) per 10 μl siRNA (final concentration 100 nM) in serum-free medium. Cells were rinsed with medium after 16 h incubation and cultured for a further 56 h before analysis. Control transfection was also performed using oligofectamine without nucleic acids (Supplementary Fig. 2 and data not shown).

Subcellular fractionation

A total of 10⁷ cells were resuspended with five volumes of buffer A (20 mM HEPES-KOH, pH 7.5, 10 mM KCl, 1.5 mM MgCl₂, 1 mM Na₂-EDTA, 1 mM Na₂-EGTA, 1 mM dithiothreitol, and 0.1 mM phenylmethyl sulphonyl fluoride) containing 250 mM sucrose. Cells were homogenized with 25 strokes of a Teflon homogenizer, and centrifuged two times at 750g for 10 min at 4 °C. Supernatants were centrifuged at 10,000g for 20 min at 4 °C. The resulting mitochondria-containing pellets were washed twice with buffer A, then resuspended in buffer A containing 250 mM sucrose. The supernatants of the 10,000g spin were further centrifuged at 100,000g for 1 h at 4 °C to produce cytosol.

Cytochrome c release assays

Mitochondria were isolated from HCT116 cells by differential centrifugation as above. Purified recombinant human Bax protein⁸ (400 ng for each sample) was pre-incubated with or without 100 μM synthetic HN peptide or mutant HN(C8P) peptide for 10 min at 4 °C. The untreated or peptide pre-treated Bax protein samples were then mixed with an equal amount of HCT116 mitochondria (in a volume of 40 μl) at 30 °C for 3 h. Samples were then centrifuged at 10,000g for 20 min to obtain pellet and supernatant fractions, measuring cytochrome c by immunoblotting.

Received 25 February; accepted 31 March 2003; doi:10.1038/nature01627.

Published online 4 May 2003.

1. Ranger, A. M., Malynn, B. A. & Korsmeyer, S. J. Mouse models of cell death. *Nature Genet.* **28**, 113–118 (2001).
2. Wolter, K. G. *et al.* Movement of bax from the cytosol to mitochondria during apoptosis. *J. Cell Biol.* **139**, 1281–1292 (1997).
3. Hashimoto, Y. *et al.* A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and Aβ. *Proc. Natl Acad. Sci. USA* **98**, 6336–6341 (2001).
4. Hashimoto, Y. *et al.* Detailed characterization of neuroprotection by a rescue factor humanin against various Alzheimer's disease-relevant insults. *J. Neurosci.* **21**, 9235–9245 (2001).
5. Nechushtan, A., Smith, C., Hsu, Y.-T. & Youle, R. Conformation of the Bax C-terminus regulates subcellular location and cell death. *EMBO J.* **18**, 2330–2341 (1999).
6. Fesik, S. W. Insights into programmed cell death through structural biology. *Cell* **103**, 273–282 (2000).
7. Tajima, H. *et al.* Evidence for *in vivo* production of Humanin peptide, a neuroprotective factor against Alzheimer's disease-related insults. *Neurosci. Lett.* **324**, 227–231 (2002).
8. Suzuki, M., Youle, R. J. & Tjandra, N. Structure of bax: Coregulation of dimer formation and intracellular localization. *Cell* **103**, 645–654 (2000).
9. Jin, C. & Reed, J. C. Yeast and apoptosis. *Nature. Rev. Mol. Cell Biol.* **3**, 453–459 (2002).
10. Xu, Q., Ke, N., Matsuyama, S. & Reed, J. C. in *Methods in Enzymology* (ed. Reed, J. C.) 283–296 (Academic, San Diego, 2000).
11. Wei, M. C. *et al.* Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death. *Science* **292**, 727–730 (2001).
12. Kermer, P. *et al.* BAG1 is a regulator and marker of neuronal differentiation. *Cell Death Differ.* **9**, 405–413 (2002).
13. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
14. Zhang, L., Yu, J., Park, B. H., Kinzler, K. W. & Vogelstein, B. Role of BAX in the apoptotic response to anticancer agents. *Science* **290**, 989–992 (2000).
15. Nouraini, S., Six, E., Matsuyama, S., Krajewski, S. & Reed, J. C. The putative pore forming-domain of bax regulates mitochondrial localization and interaction with bcl-X_L. *Mol. Cell Biol.* **20**, 1604–1615 (2000).
16. Zelphati, O. *et al.* Intracellular delivery of proteins with a new lipid-mediated delivery system. *J. Biol. Chem.* **276**, 35103–35110 (2001).

17. Jurgensmeier, J. M. *et al.* Bax directly induces release of cytochrome c from isolated mitochondria. *Proc. Natl Acad. Sci. USA* **95**, 4997–5002 (1998).
18. Stryer, L. *Biochemistry* (Freeman, New York, 1988).
19. Gray, M. W., Burger, G. & Lang, B. F. The origin and early evolution of mitochondria. *Genome Biol. Rev.* **2**, 10–18 (2001).
20. Schendel, S., Montal, M. & Reed, J. C. Bcl-2 family proteins as ion-channels. *Cell Death Differ.* **5**, 372–380 (1998).
21. Stroud, R. M., Reiling, K., Wiener, M. & Freymann, D. Ion-channel-forming colicins. *Curr. Opin. Struct. Biol.* **8**, 525–533 (1998).
22. Guo, B., Godzik, A. & Reed, J. C. Bcl-G, a novel pro-apoptotic member of the Bcl-2 family. *J. Biol. Chem.* **276**, 2780–2785 (2001).
23. Atherton, E. & Sheppard, R. C. *Solid-phase Synthesis* (Oxford Publishing, New York, 1989).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We wish to thank the NIH and the Department of Defense for its generous support. We also thank R. Cornell and A. Sawyer for manuscript preparation, and B. Vogelstein, D. Bredeesen, R. Youle and I. Nishimoto for reagents.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C.R. (jreed@burnham.org).

corrigenda

Fungus-growing ants use antibiotic-producing bacteria to control garden parasites

C. R. Currie, J. A. Scott, R. C. Summerbell & D. Malloch

Nature **398**, 701–704 (1999).

We reported in this Letter that, on the basis of its cell-wall chemistry, the bacterium associated with the fungus-growing ant *Acromyrmex octospinosus* is in the genus *Streptomyces* (Streptomycetaceae: Actinomycetes). It has been brought to our attention by *Nature* that R. Wirth, T. Wagner, C. Kost, I. Böttcher, W.-R. Arendholz and M. Redenbach (manuscript submitted) do not find evidence of a specialized relationship between bacteria in the genus *Streptomyces* and fungus-growing ants in the genus *Acromyrmex*. Our ongoing molecular phylogenetic analyses reveal that the specialized symbiotic bacterium associated with *Acromyrmex* is not a species of *Streptomyces*, but is instead in the actinomycetous family Pseudonocardaceae (C.R.C. and M. Cafaro, manuscript in preparation). This genus-level misidentification does not affect our other conclusions. □

High brightness electron beam from a multi-walled carbon nanotube

Niels de Jonge, Yann Lamy, Koen Schoots & Tjerk H. Oosterkamp

Nature **420**, 393–395 (2002).

The small round spot visible in Fig. 3 does not represent the actual emission pattern, but is an artefact caused by a low-operation voltage of the micro-channel plate. This measurement error has no effect on the value of the reduced brightness as it was not determined from the measurement of the emission pattern. □

addendum

HIV-1 superinfection despite broad CD8⁺ T-cell responses containing replication of the primary virus

Marcus Altfeld, Todd M. Allen, Xu G. Yu, Mary N. Johnston, Deepak Agrawal, Bette T. Korber, David C. Montefiori, David H. O'Connor, Ben T. Davis, Paul K. Lee, Erica L. Maier, Jason Harlow, Philip J. R. Goulder, Christian Brander, Eric S. Rosenberg & Bruce D. Walker

Nature **420**, 434–439 (2002).

The partial length HIV consensus sequences for virus A (day 18) and virus B (day 1,170) have been submitted to GenBank as accession numbers AY247251 and AY268493, respectively. □

erratum

Subsecond dopamine release promotes cocaine seeking

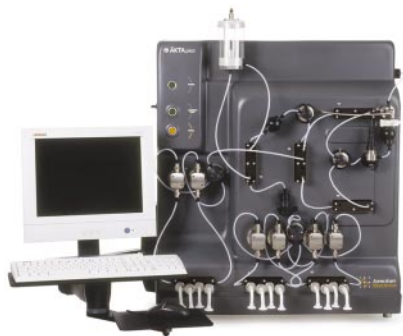
Paul E. M. Phillips, Garret D. Stuber, Michael L. A. V. Heien R. Mark Wightman & Regina M. Carelli

Nature **422**, 614–618 (2003).

In this Letter, the x axis of Fig. 4b should have ranged from –60 s to +60 s with 0 s at the grey triangle. □

Proteins pure and applied

Protein purification, scaled-up production and analysis.



Protein aplenty: AKTApilot for scale-up.

AKTApilot

Amersham Biosciences

www.amershambiosciences.com

The missing link in protein purification

AKTApilot is a bench-top separations system designed to link bench-scale process development with large-scale protein manufacturing. The system offers a range of flow rates, from 4 to 800 ml per min, and allows for multiple wavelength detection as well as pH and conductivity detection. Eight outlet valves handle fractionation. The system can accommodate an optional Frac-950 and an additional monitor for special detection requirements. It is suitable for scientists who need to perform process development and small-scale production on the same unit.

MessageSensor

Ambion

www.ambion.com

Sensitive RT

The MessageSensor reverse transcriptase (RT) kit contains an RNase H⁺ RT which, according to Ambion, results in greater sensitivity in one-step qRT-PCR experiments than RNase H⁻ RTs. It will report mRNA abundance over a 10⁶-fold range of input RNA. Transcripts can be quantified from as little as 500 fg of total RNA — 1/20th of a single cell. The RT step is completed in 15 minutes. The kit is compatible with one-step, real-time detection methods using TaqMan probes and SYBR Green.

Cryojet HT

Oxford Instruments (Superconductivity)

www.oxford-instruments.com

Broad range of temperatures for X-ray crystallography

Crystallographers can investigate molecular structures across a wide range of temperatures (90–490 K) with the Cryojet HT nitrogen jet. The device contains no moving parts: temperature and gas flow settings are adjusted using

an electronic controller unit. ObjectBench software enables users to control system settings remotely via computer and perform long-term system monitoring. The storage dewar is at atmospheric pressure, permitting liquid nitrogen refills during use. Temperature remains stable to ± 0.1 K, and the system cools down to base temperature rapidly.

Mini-UniPrep

Whatman

www.whatman.co.uk

Three tools in one

The Whatman Mini-UniPrep is a syringeless filter for removal of particulate from small sample volumes. It combines a syringe, filter and vial in one unit. Proteins can be removed from plasma, serum, whole blood and other biological fluids using the protein precipitation protocol. The method uses acetonitrile precipitation and filtration by compression to remove the protein and, according to Whatman, is three times faster than conventional spin clarification methods and removes over 99% of protein from plasma samples. It is suitable as an alternative method of removing unwanted protein prior to HPLC/MS analysis in analytical characterization in drug research.

Protein extraction/labelling

BD Biosciences Clontech www.clontech.com

One-step extraction of total cellular protein

The BD Clontech protein extraction and labelling kit provides a gentle, non-denaturing method for preparing a total protein extract of a biological sample. The extraction protocol is fully compatible with a variety of downstream analyses, including two-dimensional polyacrylamide gel electrophoresis, mass spectrometry and immunoassays such as western blots. The kit is also suitable for preparing samples to be used with BD Clontech's Ab microarray, and can be optimized for any mammalian tissue or cell type.

ProteoProfile

Sigma-Aldrich

www.sigma-aldrich.com

In-gel N-deglycosylation kit

The ProteoProfile kit provides a convenient and reproducible method for N-deglycosylation and tryptic digestion of protein samples from 1D or 2D polyacrylamide gel slices for subsequent mass spectrometry (MS) or high-performance liquid chromatography analysis. Proteomics-grade PNGase F is included to specifically cleave N-linked carbohydrates while leaving the peptide structure intact. Proteomics-grade trypsin improves digestion and simplifies MS analysis. The kit works with gel

slices or plugs and is suitable for Coomassie Blue and colloidal Coomassie stained gels. It can also be used with gels that have been silver stained and then destained with the company's ProtetoSilver Plus silver staining kit.

Pholipidec

Advanced Separation Technologies

www.astecusa.com

Column for separation of key phospholipids

Pholipidec is a stationary phase that can quantify seven key phospholipids present in soy lecithin preparations. It is tailored to account for other contaminants that may be present. The column is suitable for separation of phosphatidylethanolamine, phosphatidylinositol, phosphatidyl serine, phosphatidylcholine (lecithin), phosphatidic acid, lysophosphatidylcholine and lysophosphatidylethanolamine. Column size is 250 $\times$ 4.6 mm, including the integral guard column.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Senior European Sales Manager:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/

Iberia/ Belgium:

Evelina Rubio Hakansson (4973)

Scandinavia: Silke Opstrup (4994)

France/ Switzerland:

Amelie Pequignot (4974)

Production Manager: Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria:

Patrick Phelan, Odo Wulffen

Tel +49 89 54 90 57 11/-2

Fax +49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,

10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Haraikatomachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Hideki Watanabe

e-mail: h.watanabe@naturejpn.com

naturejobs

Cracking the combination

A solid understanding of basic research coupled with developed skills in clinical medicine is a highly desirable blend. Together, these qualities can help to turn cutting-edge science into real-world treatments. But such combined skill sets remain relatively rare. Despite an increase in the number of training programmes aimed at this career path, not enough scientists are choosing to move in this direction.

That was the message for attendees of this year's Days of Molecular Medicine symposium held in San Diego recently (K. Schwartz and J.-T. Vilquin *Nature Med.* **9**, 493–495; 2003). Many of the latest initiatives were discussed at the meeting, such as the multidisciplinary programmes set up by the University of California, San Diego, to train physician-scientists, pharmacist-scientists, PhDs who work in translational medicine, and physician-pharmacists.

Other programmes include Avenir, sponsored by France's biomedical research agency, INSERM. This supports training for clinicians involved in patient-based research, as well as institutional researchers, with renewable five-year contracts. Industry was also discussed — Genentech, Abbott Laboratories and Serono are among those that have joint training programmes with academia either in the works or under way.

Academia, industry and government seem to have recognized the need for hybrid skill sets, and the development of these programmes is a positive step in the right direction. The question remains about how to entice scientists to follow these paths — which are inevitably longer and more arduous than separate trails to basic research or clinical practice. Perhaps the market can provide some help in the shape of higher salaries and more job opportunities. But governments could also do their share — offering loan forgiveness as well as more long-term support.

Paul Smaglik

Naturejobs editor



Contents

MOVERS

Fresh faces for Harvard's genetics centre; new director at international telescope; bioinformaticist returns to his roots; and more

p464

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

Transitions

X Last month Steven Ruben left one genome-based drug-discovery company in Rockville, Maryland, for another. After ten years with Human Genome Sciences, most recently as vice-president of research, he joined Celera Genomics as vice-president of protein therapeutics.

X Michel Lannoo has been appointed scientific director of the department of physical sciences and mathematics at the CNRS, the French national research agency. He succeeds **Elisabeth Giacobino**, who has been appointed director of research at the French ministry of youth, education and research.

X David Weatherall, emeritus regius professor of medicine at the University of Oxford, UK, has been named a Fogarty scholar in residence of the US National Institutes of Health.

X Michael Wolff Jensen last month resigned as chief financial officer of Copenhagen-based biotech company Genmab to take on the same position with Pharotech, a Lundbeck spin-off, also in Copenhagen.

X Frederick Rickles will join the Federation of American Societies for Experimental Biology as executive director in July. He is currently associate vice-president for health research, compliance and technology transfer, and professor of medicine and paediatrics at the George Washington University Medical Center in Washington.

X Albert Zlotnik last month left Eos Biotechnology in South San Francisco to become senior director of molecular medicine at Neurocrine Biosciences in San Diego.

BIOLOGY

A new building at Harvard Medical School has helped to lure a trio of scientists away from the Baylor College of Medicine in Texas to Cambridge, Massachusetts. Professor of biochemistry **Stephen Elledge** and his wife **Mitzi Kuroda**, a professor of cell biology and human genetics, are set to move to Harvard's Center for Genetics and Genomics, which is directed by **Raju Kucherlapati**. At the same time, **Wade Harper**, a professor of biochemistry and molecular biology at Baylor, will join Harvard's pathology department.

For Elledge and Kuroda, the move marks the second time that the couple have conquered the 'two-body problem' that can arise when partners both work in science. The couple's first joint move came in 1989 when they left their postdocs at Stanford University to head for Baylor. "Getting two jobs to begin with was difficult," says Elledge. The move was easier this time because they were both recruited simultaneously, instead of being engaged in a dual job hunt. Although they have both enjoyed their time at Baylor, new facilities and the chance to be a part of the Harvard centre proved to be an irresistible prospect for the pair.

Like Kuroda and Elledge, Harper says he was happy at Baylor. As well as the new facilities, Harvard was attractive to him because of the collaborations he has already established there — especially in proteomics. The pathology department and the medical school's general environment offer "an outstanding opportunity to pursue our analysis of genes controlling cell division and other signal transduction pathways", says Harper.

ASTRONOMY

For **Christian Veillet**, who takes over as executive director of the firm that runs the Canada–France–Hawaii Telescope (CFHT) this month, the road to his new position called for some personal adaptations. First, he had to learn to live with a series of short-term appointments away from home — he has had a string of three-year appointments as a resident astronomer at the telescope since 1996. Second, he had to adapt to a culture in which his native French isn't the dominant language.

Judging by the dearth of applications from French astronomers for his earlier post as resident astronomer at the CFHT, there is still some resistance to such drastic moves in order to foster a scientific career. But Veillet thinks that this may be changing, although senior scientists need to spread the message. "Young people need to know they have to move," he says. His own first move might have been more of a wrench had he not had



Stephen Elledge



Mitzi Kuroda



Wade Harper



George Michaels



Fred Hassan

an appointment in France to fall back on — not to mention 16 years' experience at the centre for study and research in geodynamics and astrometry in Grasse.

BIOINFORMATICS

After finishing a research programme at biotech company Monsanto in St Louis, Missouri, **George Michaels** last month returned to his academic roots as director of bioinformatics at the Pacific Northwest National Laboratory in Richland, Washington. He was attracted by the post's industrial approach to academic problems. "There was an opportunity to do new work in bioinformatics at a scale that hadn't been done at academic institutions before," he says. The lab will play a large role in characterizing the genetic sequences of many microbial organisms.

Michaels' 30-year career includes stints at the US National Institutes of Health and George Mason University in Fairfax, Virginia, where he founded one of the nation's first doctoral programmes in bioinformatics. He now wants to use his experiences at Monsanto of leading multidisciplinary teams. He also intends to take a cue from industry in making databases and software more generally applicable — for instance, by creating a tool for microarray analysis that also works for proteomics or metabolism analysis.

PHARMACEUTICALS

Fred Hassan, who last month became chairman and chief executive of New Jersey-based drug firm Schering-Plough, knew early on that he wanted to combine science and business. After he finished his degree in chemical engineering at Imperial College, London, he went back to his native Pakistan to work in the bulk-chemical industry. But while there, he realized that he needed more management training, so he went to Harvard Business School to get an MBA.

"When I went to the business school, I realized that the real long-term value would be in innovation," he says. That led him into the drug industry. He spent 17 years with Sandoz (now Novartis), eventually heading its US operation, two years with American Home Products (now Wyeth) and six years with Pharmacia & Upjohn (now Pharmacia). Perhaps his biggest realization was that he should address the present and let the future sort itself out. "I'm the kind of person who focuses very hard on the immediate job," Hassan says.

His immediate task at Schering-Plough means several challenges, including dealing with some of the company's biggest sellers becoming available as generic drugs. The answer is, of course, innovation, he says.

CONTACTING US AT MOVERS

Please send any information on appointments or job moves to:
 ◆ naturejobseditor@naturedc.com